Interactive Shadow Removal

Gong, Han

Award date:
2015

Awarding institution:
University of Bath

Interactive Shadow Removal

submitted by

Han Gong

for the degree of Doctor of Philosophy

of the

University of Bath

Department of Computer Science

February 2015

Abstract

Shadows are ubiquitous in image and video, and their removal is of interest in both Computer Vision and Graphics. In this thesis, four methods for interactive shadow removal from single images are presented. Their improvements are made in user interaction, quality and robustness of shadow removal. We also show our state-of-the-art ground truth data set with variable scene categories for shadow removal and applications for shadow editing and its extension to video data processing.

We first give an introduction of this thesis followed by a background chapter summarising properties of shadows and previous work on image shadow removal. The main-body is organised as follows:

- Texture preserving shadow removal: This shadow removal method only requires single rough strokes to indicate shadow regions using a fusion image. We show that the approach, which includes interval-variable intensity sampling, local group fitting, and sample replacement, produces more plausible shadow removal results. We also introduce simple gradual colour transfer to correct post-processing artefacts.

- Artefact-resistant shadow removal: We show a shadow removal method with further improvements based on our texture preserving shadow removal, which includes a simplified variant of single rough strokes for user interaction, intelligent sample selection and de-noising, an efficient illumination estimation, a boundary artefact remover, and a simple yet more robust colour correction. These designs improve the quality, speed, and robustness of shadow removal.

- Shadow removal data set for variable scene categories: To support a thorough and reliable quantitative evaluation, we present a categorised and quantitatively verified data set for shadow removal in variable scenes which overcomes limitations in previous data set.

- Fast shadow removal: The presented shadow removal method has significant improvements in speed and quality of shadow removal and compatibility of different types of shadows. This is achieved by fast sample alignment using penumbra unwrapping and its intensity conversion image, and a fast and adaptive estimation

of illumination using sample synthesisation.

- Multi-scale shadow removal: We present a shadow removal method based on our fast shadow removal which resolves the issues in preserving spatial smoothness and properly relighting surfaces containing various texture and colour. This is achieved by the use of multi-scale sample synthesisation for illumination estimation and a multi-scale colour correction.

- Interactive shadow editing: Based on our shadow removal algorithms, an application for shadow editing from single images is presented.

Finally, our contributions and future research directions are concluded.

Acknowledgements

I would like to thank God for giving me an opportunity to pursue a PhD with the help of two brilliant supervisors. I would like to thank my PhD supervisors, Dr. Darren Cosker and Dr. Matthew Brown, for their ideas, advice, patience and support. Without them, this work would not have been possible. I would like to thank my colleagues Dr. Hongping Cai, Dr. Chuan Li, Dr. David Pickup, Dr. Dmitry Kit, Dr. Damien Teney, Dr. Wenbin Li, Dr. Garry Ren, Dr. Elizabeth Gabe-Thomas, Dr. Gaoyang Feng, Nick Westlake, Qi Wu, Rui Tang, and Jeehang Lee. I have enjoyed the company of these brilliant researchers who were kind enough to share their invaluable advice with me. I would like to thank Dr. Adam Kinnison, Elizabeth Jewett, and Frances Lewis for their writing advice. And last but not least, I would like to thank my family and friends for their patience, understanding and support of my work.

My father shared with me his regret over not pursuing a PhD when he was a young man. I am proud to honour him with this accomplishment.

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

Introduction

Shadows are ubiquitous and their removal is of interest in both Computer Vision and Graphics research. Although shadows can be useful cues, e.g., shape from shading, they can also affect the performance of algorithms, e.g., unwanted shadow and highlight boundaries causing artefacts in image segmentation and contributing to drift issues when tracking given moving scenes. Their removal and editing is also often the painstaking task of graphical artists. Artists may need to modify the illumination in image or edit the texture of surface but keep the original illumination. A successful shadow removal method should seamlessly relight the shadow area while keeping the lit area unchanged. When the illumination information of shadows is extracted, the removal as well as advanced editing for illumination or texture becomes possible. A shadow removal algorithm generally consists of two stages which are shadow detection and relighting. The shadow detection process in this thesis is done semi-automatically guided by user inputs.

Shadows are the areas of a scene which are darker than their surroundings due to the obstruction of direct light. Shadows can be classified as self shadow and cast shadow. Self shadow is the shadow cast on the surface of occluder while cast shadow refers to the shadow cast on the surfaces of the other objects. A shadow is generally defined as having an umbra and penumbra area – denoted by the central shadow region and its border (penumbra) transitioning illumination between the fully dark and lit area. There are some common attributes (e.g. Fig. 1-2) of shadow that can significantly increase the difficulty of their removal:

- **Texture of cast surface** Strong texture causes higher intensity variation which makes it difficult to extract illumination change from intensity changes. In addition, the appearance of dark texture is similar to shadow, which can often confuse shadow detection algorithms.

- **Softness of shadow** Softness of shadow generally relates to the width of a shadows penumbra. Higher softness brings challenges in preserving penumbra

texture when removing shadow. When the illumination change becomes much weaker than the intensity change caused by texture, it can be difficult to extract the component of illumination change.

- **Brokenness of shadow** Broken shadows contain variable illumination attributes such as irregular shape, highly varying penumbra size, and overlapping penumbra. Fixed illumination models are incompatible with such irregular illumination changes and can lead to artefacts.

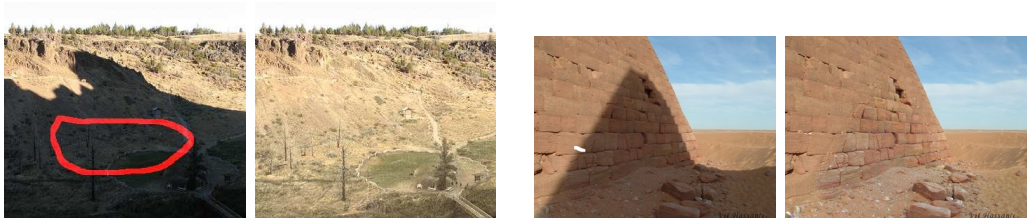
- **Colour of shadow** When shadows are not conventionally grey but colourful, it is not only difficult for machines to detect its appearance but even for humans. Also, even when shadows of this kind are detected, their removal is still difficult as the colour in the umbra could be the surfaces' reflections.

Besides these attributes, image post-processing, e.g. gamma correction and JPEG image compression, break the linearity of intensity which brings artefacts, such as inconsistency of tone and contrast, after theoretically relighting the shadow area.

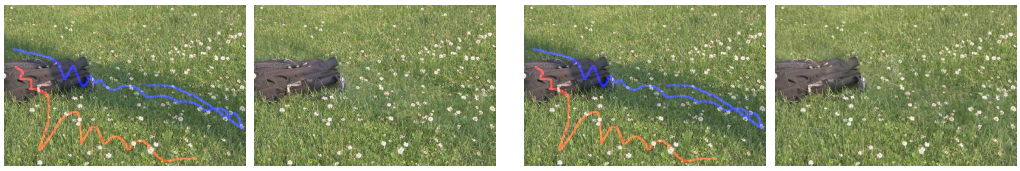
Previous state-of-the-art work has solved the problem of removing simple shadows such as compact and monotonous shadows. The problems of removing shadows from scenes containing various types of shadows as mentioned above, faithful recovery of penumbra area, and balancing robustness of shadow detection and convenience of user interaction still remain unsolved. In Computer Vision community, there is also a lack of comprehensive and reliable ground truth data set and an open benchmark for shadow removal.

In this thesis, we present four methods for user-assisted shadow removal progressively addressing these issues.

Firstly, in Chapter 3, we present our texture-preserving shadow removal method with the initial focus on simplifying user interactions, reducing artefacts in curve shadow boundaries, and producing more plausible umbra recovery. The user interactions in most previous work require users to define a precise mask of shadow region or to draw boundaries of penumbra. We introduce a new interaction that only requires a defined rough area which exhibits the characteristics of shadow (e.g. Fig. 1-1(a)). This is achieved by growing an active contour on an illumination-sensitive image generated by fusing channels from different colour spaces. To preserve penumbra texture, we sample intensity profiles perpendicular to the shadow boundary and adopt strategies for determining the density of sampling lines and the selection of good samples. To ensure the smoothness of penumbra recovery, we also introduce a local group sample fitting. Artefacts may also appear due to image post-processing. We address this using a gradual colour transfer by aligning the intensity statistics of lit and shadow areas. The evaluation at this stage is done by examining the results on some representative images due to the lack of reliable data set for shadow removal.



(a) texture preserving shadow removal (b) artefact resistant shadow removal



(c) fast shadow removal (d) multi-scale shadow removal

Figure 1-1: *User-assisted shadow removal using our four methods. (a): the red stroke covering a representative area of shadow surface is the user input (Chapter 3); (b): the white stroke covering both shadow and lit pixels is the input (Chapter 4); (c): the red stroke covers lit pixels and the blue stroke covers the shadow pixels (Chapter 6). The user input for (d) is the same as that for (c) (Chapter 7).*

In Chapter 4, we present our artefact-resistant shadow removal with further improvements in user interaction, efficiency of illumination estimation, removing artefacts in penumbra, and robustness of colour correction. This method only requires user-defined flexible single strokes covering both the shadow and lit pixels, e.g. Fig. 1-1(b), which makes its user interaction easy for complex scenes. Shadow detection is achieved by on-line training a KNN classifier (Friedman et al., 1977) for pixels measured according to user-defined samples which are automatically pre-clustered into lit and shadow halves. The speed of illumination estimation is also improved due to the simplified minimisation function and a lightweight scheme for sample selection based on the assumption of similar illumination change vector. A smoothing method for removing residual penumbra artefacts is introduced based on the assumption that neighbouring intensity samples should have similar profile of intensity change. Previous unstable colour correction of our first method is also resolved by a different approach to align the variance of intensity between shadow and lit areas using bilateral filtering (Paris and Durand, 2009).

(a) texture (b) softness (c) brokenness (d) colourfulness

Figure 1-2: *Examples of test-cases with four common shadow attributes in our data set.*

In Chapter 5, to standardise the evaluation of shadow removal, we present the first stable data set to evaluate removal of shadows in variable scene categories (e.g. Fig. 1-2). The images in our data set are carefully captured using special rigs and quantitatively verified to reject capture failures due to unavoidable natural effects, e.g., wind can cause camera shake or the sun might move behind the clouds. This data set combines 28 quantitative verified images from a previous popular data set (Guo et al., 2012) and an additional extension of 186 valid images in variable scenes captured by us. This results in a large data set which contains 214 images in total. The images are also annotated according to the four common attributes of shadow mentioned above and labelled using three intensity degrees from weak to strong. We apply an error ratio as a new measurement which reveals, to what extent, a shadow removal method has reduced the effect of shadow. This new measurement is independent of the size and darkness of shadow in each image. Our data set is released along with the first on-line shadow removal benchmark for open comparison. The evaluation on this data set reveals the weakness of our first shadow removal method as well as the previous state-of-the-art in processing complex types of shadows.

In Chapter 6, we present our fast shadow removal which shows some simple yet efficient ideas for improving robustness of shadow detection, speed of processing, and quantitatively verified compatibility of different types of shadow. This method does not assume a specific illumination change model. Although our artefact-resistant shadow removal provides the simple user input, it might be difficult for the user to choose the best strokes in practice. To address robustness and repeatability, we require users to supply two separated strokes covering samples of lit and shadow areas respectively (e.g. Fig. 1-1(c)). As for illumination estimation, the processing is done on a penumbra unwrap. We introduce a fast sample alignment based on frequency filtered intensity conversion image, and illumination change estimation using clustering.

In Chapter 7, we present our multi-scale shadow removal which improves texture details of penumbra recovery, the performance of removing shadows on texture varying surfaces and reduction of artefacts due to post-processing in the umbra area. A more accurate yet efficient sample alignment is introduced to alleviate the artefacts in the penumbra due to relighting pixels at incorrect positions. Illumination estimation by multi-scale smoothing is proposed to improve the compatibility of smoothly removing shadows on texture and colour varying surfaces. We also present a multi-scale colour correction algorithm that robustly aligns intensity variation between shadow and lit areas. We show a summary of our evaluation results for our four methods and the other competitive methods using our state-of-the-art data set. The evaluation includes the quality of shadow removal, its robustness, and the time spent on removing shadows. We also provide an analysis about the affectiveness of shadow removal quality for different shadow attributes, such as softness, brokenness, colourfulness, and texture of

background.

In Chapter 8, we demonstrate an application of our shadow removal algorithm for interactive shadow editing.

1.1 Main Contributions

The main contributions are summarised as follows:

- Texture-preserving shadow removal: We propose a method with improvements in simplifying user interaction, curve penumbra boundary recovery, and plausible umbra recovery.
- Artefact-resistant shadow removal: We propose an improved algorithm for texture-preserving shadow removal. The improvements includes simplified user interaction for complex scenes, higher processing speed, penumbra residual artefact removal using scale smoothing, and robuster colour correction using variation alignment.
- First stable and categorised data set for shadow removal evaluation: This data set contains quantitatively verified high quality ground truth data in variable scene categories along with an open on-line evaluation site.
- Fast shadow removal: We propose additional technical improvements for both speed and compatibility of variable types of shadow using sub-group illumination change synthesisation.
- Multi-scale shadow removal: We propose a method which improves smoothness of penumbra recovery on surfaces with varying texture using multi-scale smoothing and remove artefacts due to image post-processing, using multi-scale colour correction.
- Interactive shadow editing: We present a system for interactive shadow editing from single images which includes the manipulations of shape, distribution, sharpness and darkness of shadows according to the features of existing shadows.

1.2 Related Publications

The following peer-reviewed publications related to this work were produced during my PhD research:

- Gong, H. and Cosker, D. (2014b). Interactive shadow removal and ground truth for variable scene categories. *International Journal of Computer Vision (IJCV)*. under revision

- Gong, H. and Cosker, D. (2014a). Interactive shadow editing from single images. In *Workshop on User-Centred Computer Vision, Asian Conference on Computer Vision (ACCV)*

- Gong, H. and Cosker, D. (2014c). Interactive shadow removal and ground truth for variable scene categories. In *British Machine Vision Conference (BMVC)*. Best Student Paper

- Gong, H. and Cosker, D. (2014d). User-assisted image shadow removal. *Computer Vision and Image Understanding (CVIU)*. under review

- Gong, H., Cosker, D., Li, C., and Brown, M. (2013). User-aided single image shadow removal. In *IEEE Proc. International Conference on Multimedia and Expo (ICME)*

Chapter 2

Background

In this chapter, we first explain shadow formation and image formation. We then introduce the features of a shadow useful for shadow detection and the current methods for shadow detection and removal. Finally, previous methods and data sets for shadow removal evaluation are discussed.

2.1 Shadow Formation

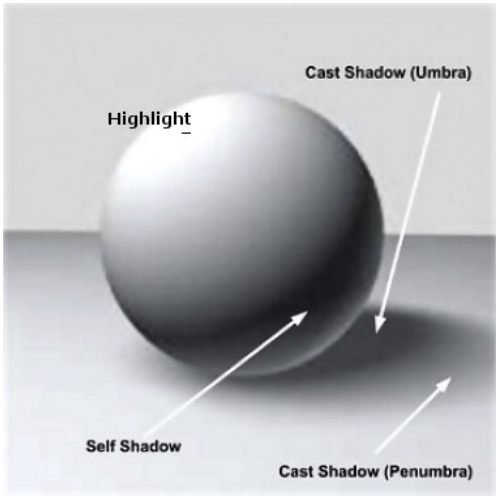


Figure 2-1: *Self shadow and cast shadow (umbra and penumbra) (Al-Najdawi et al., 2012).*

Shadows are ubiquitous in images and video data of natural scenes. The definition of a shadow is based on the difference of illumination intensity for one area and its surrounding areas. There are no absolute shadow areas. Shadows are the areas of a scene which are darker than their surroundings when direct light is blocked by an occluder or several occluders. Shadows can be categorised as either self shadows or cast shadows. As shown in Fig. 2-1, self shadow is the type of shadow cast on the occluder itself while cast shadow refers to the type of shadow cast on the objects other than

the occluder. A shadow can be divided into two parts: penumbra and umbra. Umbra is the darkest part of a shadow whilst the darkness in penumbra varies gradually as the light to penumbra is partially blocked. An illustration of the cause of a shadow is shown in Fig. 2-2.

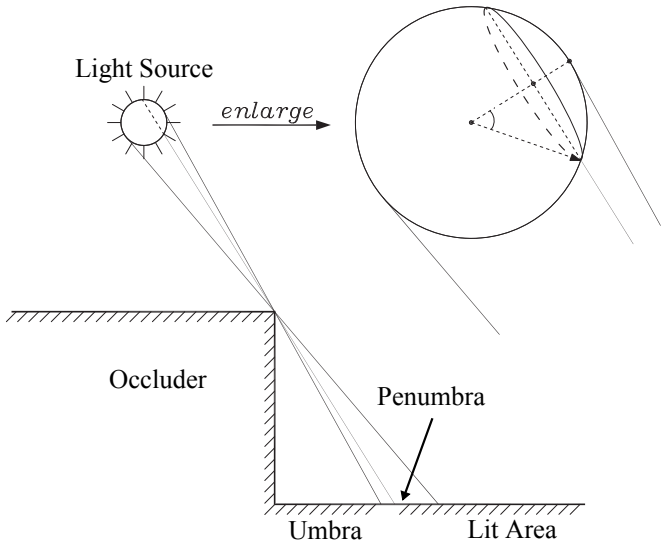


Figure 2-2: *Cause of a shadow. Umbra is the part of a shadow where direct light does not arrive at all while penumbra is the part where some direct rays are blocked. The lit area is the part where all direct rays can reach. This figure is modified from (Huang et al., 2011).*

As for the formation of penumbra in sun and sky model, Huang et al. has proofed two theorems as follows:

Theorem 2.1 (Penumbra width)**.** *The penumbra width is linearly proportional to the occluder's height and inversely proportional to the square of the cosine of the sun's incidence angle (Huang et al., 2011).*

Theorem 2.2 (Rate of illumination change)**.** *Within any image of a scene lit only by the sun and sky, if we choose any straight line perpendicular to a shadow boundary, then the illumination intensity change (or radiance) has a rate proportional to* $\sqrt{1-(2x/p-1)2}$ *along that line through the penumbra, where* x *is the distance to the umbra, and* p *is the penumbra size (Huang et al., 2011).*

A detailed measurement of illumination change of the sun and sky conditions and the proof of the two theorems can be found in (Huang et al., 2011). This sun and sky model assumes a single ball-like light source. However, in more general cases, the physics of penumbra formation can be far more complicated and unpredictable. This is because light sources have a number of different variables, including intensity, colour and direction. These factors can create a variety of different and complex shadows. The penumbra generated under these circumstances exhibit non-monotonous or overlapped intensity changes. The physics for these types of shadows are still unclear in literature.

2.2 Image Formation

The physics of image formation is closely related to the formation of a shadow in an image. In this section, camera image processing and the related Lambertian reflectance model are introduced.

2.2.1 Camera Image Processing

In the phase of raw data acquisition, the photons are collected by three different types of colour-filtered sensors for measuring redness, blueness, and greenness respectively. These three types of sensors are placed in an alternating order which form a Bayer pattern sensor array. The photons are integrated in a rectangular sensor area for a duration of exposure and converted to intensity values. This initial image of sensor response is called a raw image. The intensities recorded by the three types sensors are interpolated to form a three-layer image which is called a linear image. The intensity of a linear image pixel is proportional to the amount of photons the corresponding sensor has received for a duration of exposure.

In modern camera image processing, some post-processing steps are usually included to ensure optimum perceptual quality and storage cost. Human vision systems use white balancing to remove the colour cast in natural scenes (Land, 1971). However, due to defects in camera design or natural colour casting, e.g. bluish sky light scattering, the image recorded by camera is usually not white-balanced. The white balancing step, which is usually a linear scaling transform, is applied to adjust the colour of an image. The human vision system also senses lightness in a non-linear way which approximately follows a gamma or power function (Poynton, 2012). To compensate for this perceptual non-linearity, a gamma correction function is often applied to re-map the raw intensity values. A gamma correction function is a power-law function as shown in Eq. 2.1.

$$V_{out} = AV_{in}^{\gamma} \tag{2.1}$$

where V_{in} and V_{out} are the input and output intensities respectively, A is a constant (usually 1), γ is a parameter for encoding (when $\gamma < 1$) and de-coding (when $\gamma > 1$). As a result of gamma encoding, more quantitative intensity levels are used for displaying colour information in a dark area while less such levels are used for brighter pixels. The size of an image can be too big for normal storage and transfer over the Internet, the image is thus compressed to minimise its size without losing too much visual detail and introducing visual artefacts. The most popular technique for image compression is Joint Photographic Experts Group (JPEG) compression. The design of JPEG algorithm is based on the fact that human vision system is less sensitive to frequency response to chromaticity than luminance changes (Pennebaker and Mitchell, 1993). The steps in JPEG compression are described as follows:

- The original RGB intensities are first converted to YCrCb colour space intensities. YCrCb colour space separates the potential luminance information in Y channel and the chromaticity information in Cr and Cb channels.
- The chromaticity information, i.e. Cr and Cb intensities, are down-sampled by a factor (often 2) both horizontally and vertically.
- The down-sampled image data is passed to a block transform stage. The image is split into blocks (usually in 8-by-8 pixel size). In each block, the Discrete Cosine Transform (DCT) (Ahmed et al., 1974) is applied to obtain the frequency components of the image.
- The values of the frequency components are quantised to reduce the number of bits required for storing the image data. These quantised smaller values can be encoded using some variable bit length schemes such as Huffman code (Huffman et al., 1952) and arithmetic code (Witten et al., 1987). The degree of quantisation is adjustable according to the compression quality settings.

When the ratio of compression is too high, some artefacts can appear due to the quantisation step in JPEG compression. These artefacts include: 1) visible block-like sub-images; 2) image contrast change; 3) loss of texture.

2.2.2 Reflectance Model

Most shadow removal methods assume Lambertian reflectance and a single light source. Lambertian reflectance defines an ideal surface that diffuses light uniformly. The Lambertian model can be represented as follows:

$$\rho_c = \int_\omega E(\lambda)S(\lambda)Q_c(\lambda)d\lambda \tag{2.2}$$

where ρ is the camera response, $E(\lambda)$ is the illuminant spectral power distribution, $S(\lambda)$ is the surface reflectance function, $Q_c(\lambda)$ is the sensor sensitivities, λ is the wavelength, ω is the set of integrated wavelength, and c specifies working spectrum range/channel. e.g., red, green, blue. Assume that the camera sensitivity function is a Dirac delta function $Q_c = q_c\delta(\lambda - \lambda_c)$ where q_c is the strength of the sensor $q_c = Q_c(\lambda_c)$, Eq. 2.2 can be simplified as the following:

$$\rho_c = E(\lambda_c)S(\lambda_c)q_c \tag{2.3}$$

Reasonably assume that the surface reflectance $S(\lambda_c)$ and camera sensitivity q_c do not change over time, the change of illumination $E(\lambda_c)$ is therefore essentially a scaling effect according to Eq. 2.3. When this applies to shadow removal, shadowing changes the original illumination $E(\lambda_c)$ in the lit image and scales the original lit intensity

value. Mathematically, a shadow image I_c can be considered as a Hadamard product, i.e. element-wise multiplication, of a shadow scale layer $\mathcal{S}_c$ and a shadow-free image $\mathcal{I}_c$ (Center and Barrow, 1978) as shown in Eq. 2.4 where c is a RGB channel. The scales of the lit area are 1 and other areas' scales are between 0 and 1.

$$I_c = \mathcal{I}_c \circ \mathcal{S}_c \tag{2.4}$$

The umbra is the darkest part of the shadow whilst the penumbra is the transitional shadow boundary with a non-linear intensity change between the umbra and lit area.

2.2.3 Effects of Non-Linear Processing

Linear images are ideal for shadow removal because the intensity of a pixel is proportional to the amount of photons the sensor has received and the physics laws, e.g. Eq. 2.4 and Eq. 2.3, hold perfectly. This property can be used for recognising shadows (e.g. the derivation of invariant image (Finlayson et al., 2009)) and removing shadows (e.g. inverse scaling). However, in practice, most images are non-linearly processed into rendered images. The intensity changes in these rendered images do not perfectly comply with the laws of physics. Shadow removal artefacts can appear when some shadow removal methods still adopt the theoretical physics laws to remove shadows from rendered images. Some non-linear processing, e.g. gamma correction, are reversible but the parameters for encoding are very difficult to estimate from a single image of a natural scene. Some non-linear processing, e.g. the quantisation step in JPEG compression, is theoretically lossy and not reversible. These factors make it ill-posed to recover the linear images from their corresponding rendered images.

2.3 Features for Detecting Shadow

The features of a shadow are essential for learning how to recognise a shadow. In this section, we give a review of popular shadow features for shadow detection used in previous work.

2.3.1 Intensity

Intensity is the simplest and most widely-adopted cue for detecting shadows. The area of a shadow usually appears darker than the corresponding lit areas. This assumption is simple and weak since many simple cases can break this assumption. For instance, the black squares of a chessboard can be regarded as white squares beneath shadow. Although this unreliability exists, intensity differences between the shadow side and lit side still provide some confidence for shadow detection (Huang and Chen, 2009; Huang et al., 2011; Zhu et al., 2010).

Some methods adopt RGB intensity for per-pixel classification. Given the intensity histograms of shadow and lit areas, Wu et al. use a Gaussian Mixture Model (GMM) to approximate the PDFs of the colour statistics in the shadow and lit areas respectively. These GMMs are estimated using the Expectation-Maximization (EM) algorithm (Mitchell, 1997). Given user-specified intensity sample points, Arbel and Hel-Or detect a shadow mask by region growing using Support Vector Machine (SVM) for per-pixel classification.

2.3.2 Chromaticity

Chromaticity is a comparatively reliable feature which is, to some extend, invariant to intensity change. Many methods assume that shadow areas become darker than their surroundings but their chromaticity is preserved. Some methods choose existing colour spaces that exhibit the invariance of chromaticity such as HSV (Cucchiara et al., 2003), c1c2c3 (Salvador et al., 2004), YUV (Chen et al., 2010), and normalised RGB (Cavallaro et al., 2005). The implementations of colour space conversion are computationally inexpensive.

Intrinsic image is also important and chromaticity-related. Intrinsic image decomposition divides a single image into two components: a shading component and a reflectance component which is independent of illumination. The original single image is the product of these two components. Typical intrinsic image decomposition methods make some of the following assumptions:

- Neighbouring pixels have the same reflectance if their chromaticities are similar (Bousseau et al., 2009; Shen et al., 2011; Shen and Yeo, 2011).
- Intensity discontinuities in the luminance of an image are caused by sharp reflectance changes. The illumination change is thus only smooth (Garces et al., 2012; Rother et al., 2011; Shen et al., 2011; Shen and Yeo, 2011).
- The light source is white. In another words, the image is white balanced (Bousseau et al., 2009; Garces et al., 2012).
- The reflectance component has sparsity and only consists of a small number of colours (Barron and Malik, 2012; Rother et al., 2011; Serra et al., 2012; Shen and Yeo, 2011).
- Local reflectance variations lie in 2D subspaces of the RGB colour space (Bousseau et al., 2009).
- The scene of an image contains chromatic surfaces (Bousseau et al., 2009; Finlayson et al., 2009, 2002a,b, 2006; Fredembach and Finlayson, 2006, 2005; Yang et al., 2012).

- The image is captured by narrow-band camera sensors and under Planckian lighting (Finlayson et al., 2009, 2002a,b, 2006; Fredembach and Finlayson, 2006, 2005; Yang et al., 2012).

However, chromaticity features are sensitive to distinct illumination change, non-linear image post-processing, and complex surface texture. Although the illumination information can be extracted, the derivation of an intrinsic image may remove illumination-like surface texture from its reflectance component.

2.3.3 Texture

Texture is another reliable property which is invariant to illumination change. Some methods assume that the texture of shadow areas is similar to the texture of their corresponding lit areas. Texture descriptors, such as Texton (Martin et al., 2004), are used to build the texture features of the segments at the two sides of an image edge. Texton gives a histogram of texture feature distribution. Some measurements, such as χ^2 distance (Guo et al., 2012; Lalonde et al., 2010), are used to compare the texture similarity among the image segments. Texture correlation is a reliable method for detecting shadows since textures are invariant to various image post-processing, illumination change, and colour difference. However, texture feature building and comparison are computationally expensive.

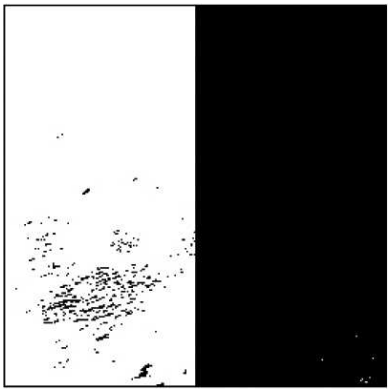
2.3.4 Physical Property

Physics based methods simplify the process of shadow detection using strict illumination assumptions and additional information for stronger evidence of a shadow's existence. They generally require non-trivial capturing devices for additional information and assume specific and constrained illumination models. These strategies provide faster performance and more accurate results but limit their applicable range of scenarios. Some popular usages are introduced in this sub-section.

Chromagenic Information

A chromagenic camera can capture two registered images of a scene: one is captured using normal camera configuration and the other through a coloured filter. An example of the image pair is shown in Fig. 2-3. Chromagenic theory has revealed that the relationship between filtered and unfiltered camera sensor responses is strongly dependant on the colour of the illuminant. Some methods, e.g. (Finlayson et al., 2007), use this property to classify different illuminants as well as detecting shadows.

(a) reflectance images from (Nascimento et al., 2002)



(b) illuminant detection using the chromagenic features (Finlayson et al., 2007)

Figure 2-3: *Left: the left and right halves of the image are illuminated by two different light source. Right: The pixels of the image are binary which indicate the index of the transform that best maps it to its filtered counterpart. Ideally, the left and right halves of the binary mask should be completely white and black without any noise points. This is because the two halves are illuminated by two different illuminants.*

Multi-Spectral Information

Multi-spectral information provides additional evidence of a shadow and is often used for shadow detection. For example, Fredembach and Süsstrunk require a special NIR/RGB camera to capture registered NIR image and RGB image of a scene (e.g. Fig. 2-4). They make the following assumptions:

- The direct light, that causes shadows to appear, is "hot" light – red and infra-red light. It stimulates significant sensor responses in the IR channel. The environmental light is almost "non-hot", e.g., blue and green, which gives insignificant sensor responses in the IR channel. Therefore, the shadow pixels consist of a relatively lower proportion of IR response.
- All objects' surfaces in a scene are non-black, i.e., with light colours.
- Shadow pixels appear darker in both NIR and RGB images.

Examples of the kinds of scenarios that are common in producing shadows are the following combinations: skylight/sunlight, fluorescent/incandescent lights, flash/incandescent lights, and flash/fluorescent lights. A plot of power distribution for these common light sources is shown in Fig. 2-5. They use a conventional brightness image and a NIR/RGB ratio image to detect shadow pixels.

Figure 2-4: *RGB (left) and NIR (right) images of a scene, captured using identical exposure settings to allow comparison. The texture of the vegetation in the lit area of the wall and the texture of the shadowed wall both appear weaker in the NIR image. The NIR image shows higher contrast for the intensities of shadow areas and lit areas. The shadowed surface appears darker in the NIR image than its corresponding pixels in the RGB image due to its lack of "hot" reflectance.*

Active Illumination

Some methods adopt active illumination which capture images of the same still scene illuminated by different light sources. Drew et al. make use of large values (outliers) in the log difference between images taken with flash and without flash (ambient) to detect a shadow mask. Yoon et al. process two images taken by switching on and off two additional light sources in turn to simulate a light source with infinite dimensions. They apply simple set operations to produce a shadow-free image without requiring any region extractions.

These methods using active illumination features require additional light sources and accurate image registration (usually compensated by short-time-interval capturing). They are suitable for restoring a shadow-free image of a single object and are not suitable for the shadow removal from trivial images.

Intensity Change

Many methods (Huang and Chen, 2009; Huang et al., 2011; Lalonde et al., 2010) analyse intensity change between potential lit and shadow samples to recognise the boundary of a shadow. Various forms of intensity change are supplied as features and the typical examples are listed as follows:

- **Ratios of intensity** This feature is a 3-by-1 vector $[\rho_r, \rho_g, \rho_b]$ which reflects the scale of intensity change. It is computed as follows:

$$\begin{cases} \rho_r = r_1/r_2 \\ \rho_g = g_1/g_2 \\ \rho_b = b_1/b_2 \end{cases} \tag{2.5}$$

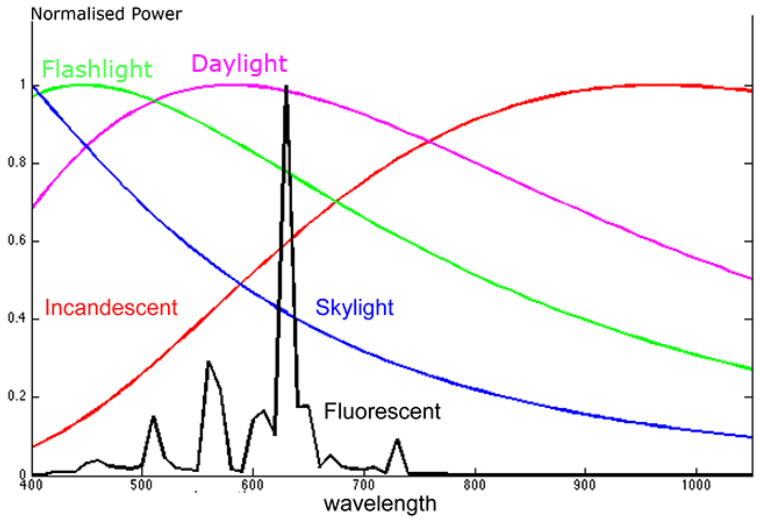


Figure 2-5: *Measured spectra of typical shadow-creating illuminants. The striking differences of sensor response appear between the NIR band (750nm-1400nm) compared to the visible band (380nm-700nm).*

(a) illumination 1 (b) illumination 2 (c) shadow removal output

Figure 2-6: *An example of image set with active illumination (Yoon et al., 2002).*

where r_1 and r_2, for instance, can be the average red intensities for two potential lit and shadow sides. As for a shadow and lit pair of the same material, the non-shadow region is assumed to have a higher intensity for all RGB channels.

- **Inter-channel ratio** Based on Eq. 2.5, a 2-by-1 feature can be computed as a vector – $[\rho_r/\rho_g, \rho_g/\rho_b]$.
- **Gradient of illumination** This feature is a 3-by-1 vector $[s_r/H_r, s_g/H_g, s_b/H_b]$ where s_r ,s_g, and s_b refer to the colour gradient magnitude at potential shadow boundaries, H_r ,H_g, and H_b refer to the RGB intensities at the potential lit side which is used to cancel the local reflectance.
- **Colour gradient direction** This feature is a 3-by-1 vector $[\gamma_r, \gamma_g, \gamma_b]$ where its elements are the angular directions of colour gradients. The valid shadow boundary points are assumed to have similar gradient directions. Outliers are the other abnormal boundaries, e.g., boundaries caused by texture or occlusion.

Fig. 2-7 shows a visualisation of this property.

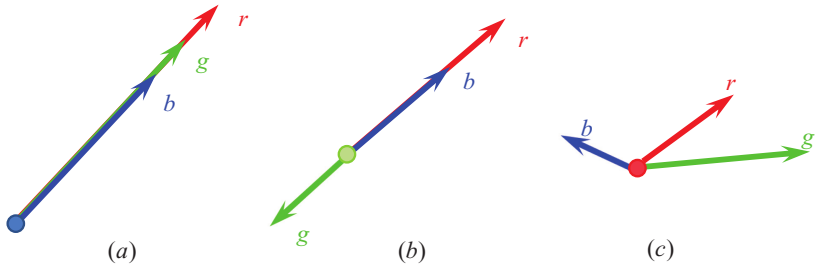


Figure 2-7: *Colour gradient vector at a shadow edge for each channel. Shadow edges (a) usually have similar gradient directions for all channels. (b) and (c) are potential edges of outliers as their gradient directions are abnormal (Huang et al., 2011).*

- **Relative colour gradient direction** This is a variant of colour gradient direction which is a 3-by-1 vector $[\gamma_{rg}, \gamma_{gb}, \gamma_{br}]$. This vector represents the angular difference for each colour gradient direction. Based on Fig. 2-7, this vector is computed as follows:

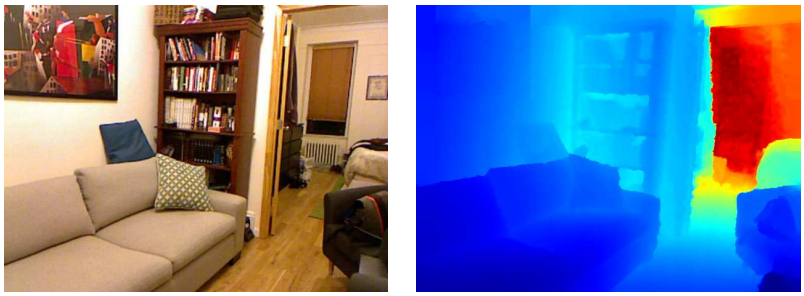
$$\begin{cases} \gamma_{rg} = min(|\gamma_r - \gamma_g|, 2\pi - |\gamma_r - \gamma_g|) \\ \gamma_{gb} = min(|\gamma_g - \gamma_b|, 2\pi - |\gamma_g - \gamma_b|) \\ \gamma_{br} = min(|\gamma_b - \gamma_r|, 2\pi - |\gamma_b - \gamma_r|) \end{cases} \tag{2.6}$$

  Empirically, this feature works better than gradient of illumination when a SVM is used as the classifier.

The red and green intensities and the ratio of red intensity to blue intensity, increase significantly from the shadow area to the lit area (Huang et al., 2011).

Geometry Information

If illuminant, geometry information, and material information are known, illumination information including shadows in a scene can be predicted. To extract shadows, some methods (Chang et al., 2002; Yoneyama et al., 2003) assume that the scene contains a single light source, a flat casting surface, known rough object shape, e.g. the shape of a human. However, these assumptions are too strict and these methods are therefore unable to deal with complicated and more general scenes. Recent developments (Barron and Malik, 2013; Chen and Koltun, 2013; Lee et al., 2012; Xiao et al., 2014) use RGB-D (Normal RGB image and its depth map) to derive intrinsics images (e.g. Fig. 2-8). Intrinsic image consists of two components: one for reflectance and another for shading (illumination). Xiao et al. use intrinsic images from RGB-D data to automatically



(a) input RGB image (b) input depth image

(c) output reflectance (d) output shading

Figure 2-8: *An example of intrinsic image decomposition using RGB-D data (Chen and Koltun, 2013).*

detect shadows. They make use of surface normals derived from depth image and assume pixels with similar normals, spatial locations and chromaticity should have similar colours. However, depth information obtained from commercial off-the-shelf products, e.g., Kinect, is often noisy which leads to unreliable shadow detection. Current popular devices use IR light and sensors for detecting depth. This limits not only its working environment to indoors but also its detectable range of depth.

2.4 Using Shadow Features for Detection

In the previous section, we have reviewed the popular features for shadow recognition. In this section, we discuss how to use these features for shadow detection. There are three popular elements used for collecting and detecting shadows:

- **Pixels** Many types of features are pixel based such as intrinsic images using geometry information (Xiao et al., 2014), active illumination (Drew et al., 2006; Yoon et al., 2002), multi-spectral information (Fredembach and Süsstrunk, 2010), chromaticity-related colour spaces (Salvador et al., 2004). As pixel-wise analysis may produce noisy shadow detection results, image filtering is usually performed to compensate for this.
- **Image edges** The discontinuity of illumination creates some edges in an image.

Shadow features lie in the local areas of two sides of each edge, such as difference of intensity, chromaticity, and texture. Some features also lie at the edges, such as colour gradient and edge width. Some intrinsic image based methods use the difference between the edge maps of the original image and its invariant image to obtain the shadow edge map. The invariant image can be a grey-scale invariant image (Finlayson et al., 2006; Fredembach and Finlayson, 2006) or an invariant chromaticity image (Finlayson et al., 2009; Fredembach and Finlayson, 2005).

- **Image segments** It is reliable to detect illumination when the surfaces on the two sides of an edge are made of the same material. When it is not the case, illumination analysis on edges may be unreliable. Guo et al. therefore adopt segments as the basic element for analysis and establish pair-wise connections among the segments based on their features. This makes it possible to process shadow surfaces and their non-neighbouring lit surfaces. Segment-based analysis can also provide less noisy results, e.g. (Finlayson et al., 2007), compared with pixel-based analysis.

Using the edge of an image is the most popular way to detect shadow whilst some others opt for the correlation among segments of an image and pixel-wise classification. The selection of edge or segment in each feature relies on the pre-processing of edge detection. Some shadow edges or segments can not be recognised if the early-stage edge detection fails.

2.5 Shadow Relighting Approaches

Given a detected shadow mask or shadow edges, the next stage of shadow removal is to carefully relight the shadow area. The shadow removal process is based on the image formation in Eq. 2.4. Due to image post-processing discussed previously in Section 2.2.1, the textures in shadowed surfaces appear weaker than its lit area's which causes the inconsistency of texture contrast. As the intensities of a shadow are not constant, especially in the penumbra, the estimation of a non-uniform shadow scale field is difficult.

2.5.1 Gradient Reintegration

Shadow images can be converted to the log domain and the multiplicative calculation in Eq. 2.4 becomes an addition as follows:

$$log(I_c) = log(\mathcal{I}_c) + log(\mathcal{S}_c) \tag{2.7}$$

The gradient of the log image represents the change of shadow scale. As the change of shadow scale is most predominant at a shadow boundary, i.e. penumbra, suppressing

the scale changes at the shadow boundary is equivalent to removing shadow effects. The shadow-free log image can be restored by reintegrating the modified gradient of the log image. There have been two ways in previous work for suppressing the strong gradient at the shadow boundary:

- **Zeroing gradient** A simple way for suppression is to set all gradient values of penumbra pixels to 0 (Finlayson et al., 2002a,b). This works fine for very hard shadow boundaries. However, this results in unrealistic smoothing in penumbra and the texture information there is lost.

- **Iterative diffusion** An iterative diffusion process is applied to smoothly fill in the derivatives in penumbra regions according to their neighbouring lit pixels (Finlayson et al., 2009, 2006). It also includes a step at each iteration to enforce integrability (Frankot and Chellappa, 1988). This method can avoid unnatural smoothing due to zeroing gradient, to some extent, but can still cause texture loss in penumbra.

There have been two ways for reintegrating a shadow-free log image from the modified gradient image:

- **2D reintegration** Given a detected shadow edge mask, some methods (Finlayson et al., 2009, 2002a,b, 2006) grow simple gradient based edges across the shadow edge mask. There are two steps in the mask growing process: (1) It iteratively dilates the mask and replaces unknown derivative values by the mean of known values; (2) It forms a second derivative, goes to Fourier space, divides by the Laplacian operator transform, and goes back to x,y space. Neumann boundary conditions bring an unknown additive constant in each recovered log colour. To solve this, for each pixel, the brightest pixels in the recovered image are mapped to the corresponding pixels in the original image. This works well for clear shadow edges, but produces artefacts at soft shadow boundaries.

- **1D reintegration** Fredembach and Finlayson use a random Hamiltonian path along which the gradient of the log image is reintegrated in a 1D manner. The Hamiltonian path enters and exits the shadow regions once. To reduce artefacts, open shadow boundaries are made closed and the number of crossings through the shadow edges are also limited. The pixels at shadow edges, which are not visited by the Hamiltonian path, are in-painted. This method is fast but has difficulty processing complex shadow scenes, e.g. broken shadow.

These methods rely on the detection of the shadow edge and work well when the shadow edges are clear and hard. However, on some trivial scenarios where shadows are soft and complex in shape, edge detection will fail. For soft shadow edges, both in-painting and gradient zeroing can cause noticeable artefacts.

2.5.2 Constant Scale Relighting

Fredembach and Finlayson assume that the shadow scales in a single umbra region are constant, which simplifies the estimation of shadow scale field. They find the constant by multiplying it to the values of pixels immediately adjacent to the shadow side of the shadow edge and minimise the difference between the multiplied values and the values of pixels immediately adjacent to the lit side of the shadow edge. Some additional strategies, such as chromaticity and intensity constraints introduced in §2.3, can be applied to increase the robustness of scale estimation which excludes implausible values. For multiple shadow regions in images, the algorithm computes an independent constant for each region. As the constants do not take account of the gradually changing illumination profile in penumbra regions, image in-painting (Criminisi et al., 2004), is adopted to compensate for its incapability of estimating varying penumbra scales. This method is fast and can produce satisfying shadow removal results for hard shadows.

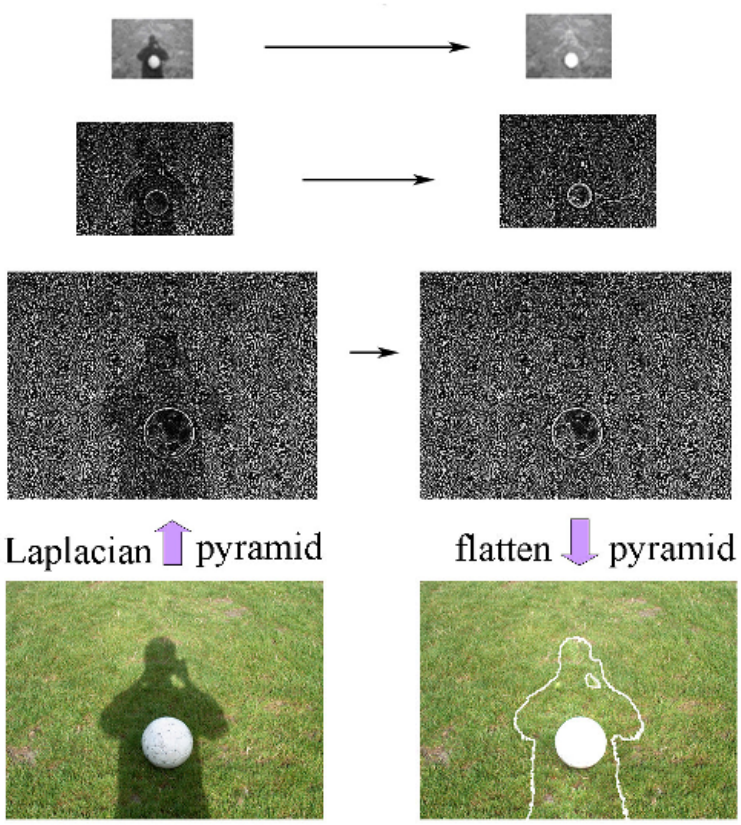


Figure 2-9: *Shadow relighting using the Laplacian pyramid. In each pyramid level, an affine model is applied for aligning the statistics in shadow area and lit area. The relit image is computed by flattening the modified Laplacian pyramid.*

Shor and Lischinski adopt a similar approach which is improved by using a multi-scale image pyramid. They make use of an affine relationship between the lit and shadowed intensities on the same surface and estimate a single shadow scale α_k as

follows:

$$\alpha_k = \mu_k(L) - \frac{\sigma(L)}{\sigma(S)}\mu_k(S) \tag{2.8}$$

where k indicate a RGB channel, L and S refer to the sets of lit and shadowed pixels near the penumbra area. μ and σ are the functions computing mean and standard deviation for a set of values. For better recovery of multi-scale details, e.g. texture, they compute the Laplacian pyramid of the image, estimate the scales for each level and finally flatten these pyramid layers to obtain the restored image. To compensate for the artefacts in non-uniform shadow recovery, they estimate different sets of affine parameters for different areas inside the shadowed regions. Their penumbra recovery still relies on image in-painting.

In summary, both of these methods can achieve fast processing performance due to the simplified illumination model. However, they do not preserve the texture in penumbra (due to the in-painting process) and can create significant artefacts when the shadow boundary is soft and complex or the umbra illumination is significantly varying.

2.5.3 Colour Transfer

Based on the statistics of intensity distributions, some methods (Wu and Tang, 2005; Wu et al., 2007) use a variant of colour transfer to relight shadow areas. These methods have the following two assumptions:

- The supplied sample intensities of shadow and lit regions are similar in texture and chromaticity but different in intensity.
- A good estimation for the shadow matte should be smooth inside umbra and non-shadowed regions and have larger and gradual gradients in the penumbra region.

The standard colour transfer function T (Reinhard et al., 2001) is defined as follows:

$$T(I_t(x)) = \mu_s + \frac{\sigma_t}{\sigma_s}(I_s(x) - \mu_t) \tag{2.9}$$

where I_t is the processed image, I_s is the original image, x is a pixel, μ_s and μ_t are the means of pixel values of the source set and the target set respectively, and σ_s and σ_t are the standard deviations of pixel values of the source set and the target set respectively.

There are many ways to modify the standard colour transfer function for shadow relighting. For instance, previous work uses a GMM to estimate the distributions of intensity for pixels in shadow and non-shadow areas respectively first. The correspondences of these Gaussians can be established by aligning them by order. Fig. 2-10 shows an example of the intensity distribution alignment between the intensities of the

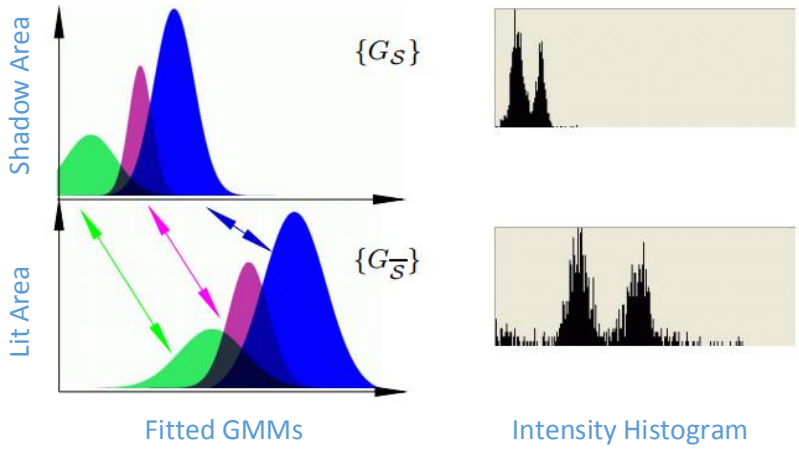


Figure 2-10: *Intensity distribution alignment between corresponding GMMs of intensity distribution.*

shadow area and the intensities of the lit area. For each pixel in the shadow area, the affinity of the pixel's colour to a particular Gaussian model is measured. A variant of colour transfer $\hat{T}$ is then performed between these Gaussians in a weighted manner as follows:

$$\hat{T}(I(x)) = \frac{1}{\sum_i d^i(x)} \sum_i d^i(x) T_i(I(x)) \tag{2.10}$$

where i is the index of Gaussian distribution, d^i is a model affinity measurement function, T_i is the standard colour transfer function. As direct colour transfer may cause artefacts at the shadow boundary, a linear distance function, which is also called spatial affinity, is used to model the gradually increasing effect of the colour transfer.

The relighting problem is formulated as an energy minimisation problem for all pixels in the image:

$$E(S_c) = E_{colour}(S_c) + E_{smooth}(S_c) \tag{2.11}$$

where $E_{smooth}(S_c)$ is the error of smoothness of the estimated S_c shadow matte, and $E_{colour}(S_c)$ measures the difference between the original image and a composited shadow image. The composited shadow image is the element-wise product of the estimated shadow scale and its coarse shadow-free image. $E_{smooth}(S_c)$ consists of the smoothness errors of umbra area and penumbra area is as follows:

$$E_{smooth}(S_c) = E^{S_c}_{smooth}(S_c) + E^{\omega}_{smooth}(S_c) \tag{2.12}$$

where $E^{S_c}_{smooth}(S_c)$ refers to the smoothness of the whole shadow matte, $E^{\omega}_{smooth}(S_c)$ refers to the smoothness at the shadow boundary. Higher smoothness is enforced at the shadow boundary to avoid artefacts.

For each pixel in the shadow area, the relighting is performed iteratively in two steps:

- Apply colour transfer to shadowed pixels in the image which takes account of both its model affinity and spatial affinity.
- Compute the energy function $E(S_c)$ and decide the next parameters for optimisation until the error is small enough.

This method preservers the texture of a shadowed surface and produces satisfying results for common shadows as well as broken shadows. As the iterative optimisation is pixel-wise, this method is computationally costly.

2.5.4 Matting

Guo et al. formulate the shadow scale field estimation as an alpha matting problem. An image I can be decomposed into a foreground layer F and a background layer B as follows:

$$I_i = \gamma_i F_i + (1 - \gamma_i) B_i \tag{2.13}$$

where i is a pixel index, γ is a scale matte. By modelling the shadow-free image as foreground and the shadow image as background, Eq. 2.13 can be rewritten as follows:

$$I_i = \gamma_i (L_d R_i + L_e R_i) + (1 - \gamma_i) L_e R_i \tag{2.14}$$

where L_d and L_e are 1-by-3 vectors representing the intensity of the direct light and environmental light, R is the surface reflectance. This matte γ can be solved by using a spectral matting algorithm (Levin et al., 2008). The shadow-free image can be computed as follows:

$$\begin{aligned} I_i^{shadow_free} &= (L_d + L_e) R_i \\ &= (\gamma_i L_d + L_e) R_i \frac{L_d + L_e}{\gamma_i L_d + L_e} \\ &= \frac{r+1}{\gamma_i r + 1} I_i \end{aligned} \tag{2.15}$$

where $r = L_d / L_e$.

2.5.5 Data Fitting

Data fitting is a popular method for shadow relighting as it preserves surface texture at the shadow boundary. These methods adopt sampling lines passing through shadow boundaries to collect intensity profiles for fitting. There have been two ways of intensity sampling:

- **Sampling along normals of shadow boundary** Some methods (Arbel and Hel-Or, 2007; Mohan et al., 2007; Su and Chen, 2010) sample intensity profiles along the sampling lines perpendicular to shadow boundaries. The normals of shadow boundaries are required for determining sampling direction at each boundary point. The computed normals can usually be noisy. (Arbel and Hel-Or, 2007; Mohan et al., 2007) require accurate shadow boundaries and (Su and Chen, 2010) allows rough shadow boundaries and the algorithm aligns the sample profiles after their collection.

- **Sampling horizontally and vertically** Liu and Gleicher propose a method to sample intensity profiles along horizontal and vertical sampling lines passing through each shadow boundary point.

The length of sampling line and the space interval between sampling lines are fixed and pre-defined. This fixed setting can cause shadow removal artefacts when the curvature and width of shadow boundaries are varying. The fitting in previous work can be categorised into curve fitting and surface fitting:

- **Curve Fitting** Curve fitting fits each sampled intensity profile to a pre-defined illumination change function. Mohan et al. model the gradient of a single intensity profile as a convolution of an impulse function whose amplitude is the intensity difference between lit and shadow sides. There is also a triangle shaped sharpness filter of width that matches the width of the penumbra. Liu and Gleicher use a C1 continuous piece-wise curve to model the intensity change in log domain. The curve has two constant segments at both ends for definite lit and shadow areas, and a segment of smooth cubic curve in the middle to approximate the gradual change of intensity of penumbra. They have also added a smoothness term in their energy function to ensure spatial smoothness of estimated shadow scales.

- **Surface Fitting** Arbel and Hel-Or assume that the shadow is cast on a surface which may be approximated by a thin plate spline (Duchon, 1977) and use multiple cubic smoothing splines to fit the surface in shadow areas. They add a smoothing term in its surface fitting function to improve spatial smoothness. A directional smoothing is also applied to suppress the band-like artefact at the shadow boundary. This approach takes account of the varying width of the penumbra and is compatible with shadows cast on curved surfaces. It also preserves texture across shadow boundaries. Its drawbacks are its slow processing speed and the fact that the algorithm has to perform the surface fitting once for every single shadow region.

2.5.6 Optimisation Using Gradient Priors

Given the sampled intensity profiles which are in same length and are perpendicular to shadow boundary, Su and Chen align these profiles according to their gradient changes and register each profile to a unique column of array. They estimate the illumination change based on two priors of this array:

- The y-direction, i.e., row-wise, gradient is mainly caused by the illumination change. The illumination change is assumed to be a smooth curve. This assumption is used to estimate the illumination change.
- The x-direction, i.e., column-wise, gradient is almost constant for each column. This assumption ensures the smoothness of illumination change in penumbra.

According to these two priors, an energy minimisation function is proposed as follows:

$$c^* = \arg\min_c \sum_{p \in R} |\nabla_y c(p) - \nabla_y I^{'}(p)| + \lambda_y |\Delta_y c(p)|^2 + \lambda_x |\Delta_x c(p)|^2 \tag{2.16}$$

where c is the shadow scale field of the penumbra array, R is the set of index of an array column, $I^{'}$ is the original intensity data, ∇_y refers to the y direction gradient, Δ_y refers to the y direction Laplacian operator, λ_x and λ_y are two weights. In this function, the first term measures the fitness of data, the second and third terms measure the estimated smoothness in y and x directions respectively. The authors also suggest a larger value for λ_x.

2.5.7 Bilateral Filtering

Yang et al. propose an approach without the requirement of shadow detection. A 2D intrinsic image is first estimated from a RGB image according to some colour cues, especially chromaticity. To estimate a 3D intrinsic image, a bilateral filter is used to enhance a target image using the texture of a guidance image. Joint bilateral filter can combine the high frequencies from one image and the low frequencies from another image. This property is very useful for transferring the details from the 2D intrinsic image to the original RGB image. With this method, a coarse estimation of the maximum chromaticity values of the intrinsic image is computed and filtered using joint bilateral filtering with the 2D intrinsic image as the guidance image. This assumes that the colours of a surface are locally similar. The filtered values are finally used to compute a 3D intrinsic image which contains higher distinction of chromaticity values of different pixels.

2.6 Colour Correction

Most images we have today are pre-processed. The linear relation between intensities of an image and the amount of photons received by a sensor does not hold precisely. These pre-processing steps can include gamma correction, image compression, and contrast adjustment which make the relation non-linear. Naive relighting can cause the inconsistent perceptual appearances between the relit regions and the original lit regions. Previous work adopt the following two approaches to compensate this artefact:

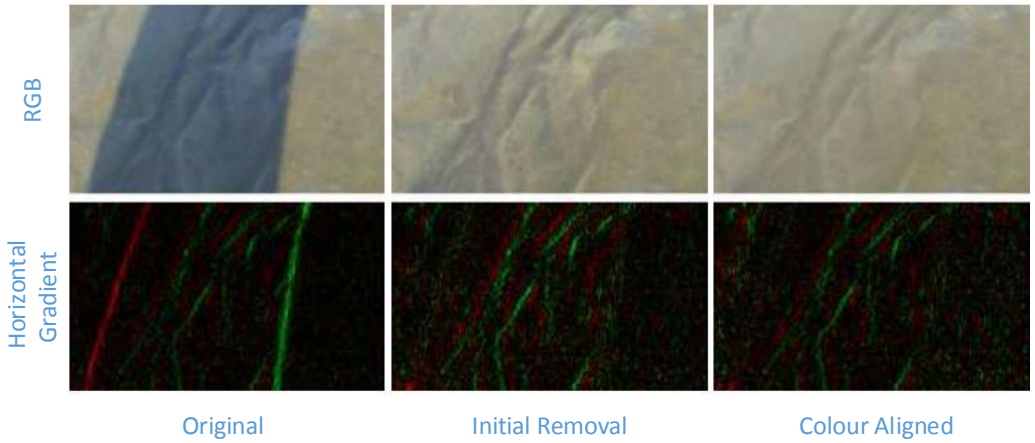


Figure 2-11: *Colour correction using gradient alignment: the texture inconsistency in the shadowed-recovered area is reduced after the gradient alignment.*

- **Gradient alignment** Liu and Gleicher apply a variant of colour transfer described in (Reinhard et al., 2001) to align the variation and mean of the relit intensity gradients between shadow and lit regions. After the alignment, a 2D reintegration is performed to restore the amended shadow-free image. Fig. 2-11 shows an example of gradient alignment.

- **Histogram specification** Arbel and Hel-Or apply histogram specification to align the colour statistics for two matching patches which are a shadow patch and its corresponding lit patch. This method is not colour-preserving, e.g., it may change the colour green to yellow. Fig. 2-12 shows an example of histogram specification.

Both of these approaches assume that the mean values can be correctly estimated by the shadow relighting algorithm and the error exists only because of the difference of variance.

Figure 2-12: *Colour correction using histogram specification: the texture inconsistency in the shadowed-recovered area is reduced after the histogram specification. Some green grass at the bottom are mistakenly changed to yellow.*

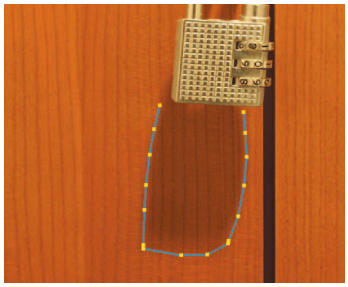
2.7 User Interaction

As we have already seen, there are some existing methods for automated shadow removal such as (Guo et al., 2012; Lalonde et al., 2010). Some of these methods rely on feature-based training and their results are dependent on the training data set and initial image segmentation. Some other automated methods (Finlayson et al., 2007, 2002a,b, 2006; Fredembach and Finlayson, 2005) make use of the intrinsic image decomposition. The decomposition process itself may not be reliable and its initial edge detection can affect the results and limit its usage to removing only hard shadows. Automated methods can often omit important shadow regions or classify other dark non-shadow regions as shadows. Compared with automated methods, user-aided methods are generally more reliable in shadow detection. The complexity of required user input varies for different methods. Fig. 2-13 shows some representative user interactions and we describe them as follows:

- **Shadow mask** Arbel and Hel-Or require users to supply a precise binary mask. This task is very time-consuming as users need to identify every shadow pixel in the image.

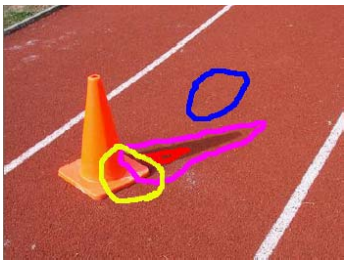
- **Sparse boundary points** Mohan et al. (Fig. 2-13(a)) require users to specify some sparse shadow boundary points. These boundary points are connected by straight lines to mark the shadow boundary. With this design, there are difficulties in marking highly varying shadow boundaries.

- **Precise shadow boundary** Liu and Gleicher (Fig. 2-13(b)) require users to draw the precise shadow boundary. The width of the shadow boundary is fixed.

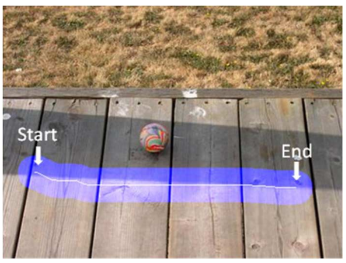


(a) sparse boundary points (Mohan et al., 2007)

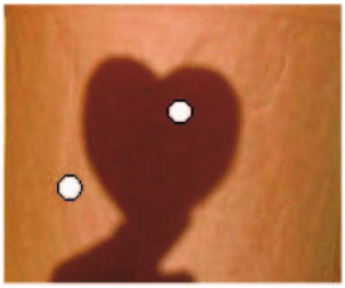
(b) precise shadow boundary (Liu and Gleicher, 2008)



(c) quad map (Wu and Tang, 2005; Wu et al., 2007)



(d) rough shadow boundary (Su and Chen, 2010)



(e) sample points (white) (Arbel and Hel-Or, 2011)

(f) single clicks (red) (Shor and Lischinski, 2008)

Figure 2-13: *Popular types of user interactions for shadow removal.*

The result is dependent on how accurate the marked boundaries are. This can take considerable time for users, when the shadow boundary is highly variable.

- **Rough shadow boundary** Su and Chen (Fig. 2-13(d)) allow users to draw rough shadow boundaries. They perform sample alignment to correct inappropriate user strokes. The boundary width can also be adjusted while drawing. However, the user can still find it difficult to mark highly varying boundaries, e.g. broken penumbra.

- **Quad map** For each single shadow region, Wu et al. (Fig. 2-13(c)) require users to specify samples of four regions which are shadow regions, non-shadow regions, uncertain regions, and excluded regions. This type of user input is complex and would be unsuitable for non-experienced users.

- **Sample points** Arbel and Hel-Or (Fig. 2-13(e)) require users to mark some sample points of shadow and non-shadow regions for each single region. These points are expaned to larger areas in order to detect a shadow mask. The complexity is that users have to mark every region and apply shadow removal for them one by one.

- **Single clicks** Shor and Lischinski (Fig. 2-13(f)) propose a simple user interaction – a single click of each shadow region. Its shadow detection is achieved by image

matting from a grown shadow seed. This is convenient for simple shadow scenes but is however not compatible for more complex cases. For example, the other shadowed surfaces are not surrounded by the initially detected surface, or in another case, the penumbra is too wide. It can also be inconvenient when there are too many shadow regions in the scene, e.g., broken shadow.

2.8 Evaluation and Data Set

To date, most shadow removal methods have been evaluated by visual inspection on some selected images – with only a few exceptions performing quantitative evaluation. This is in part due to a lack of high-quality, varied, and public ground truth data. Shor and Lischinski perform a quantitative test but comparison is difficult due to their data not being publicly available. Guo et al. provide the first public ground truth data set for shadow removal and perform quantitative testing. However, the difficulty of collecting such a data set is highlighted in their work, with the appearance of some global illumination changes and mis-registration between the shadow and shadow-free images being a difficult factor to control. This can make quantitative testing on such data somewhat difficult, as these errors can influence shadow removal results. Another desirable property, as yet not explored by existing data sets, is the categorisation of shadows. Such attributes are useful as different shadow types can present their own unique challenges, e.g. removal of coloured shadows (i.e. through a glass bottle) are more difficult than consistent unbroken shadows (i.e. a human silhouette). Having such categories in a ground truth would allow us to evaluate algorithm performance in a range of different scenarios.

2.9 Interactive Shadow Editing

Naive shadow editing using existing image manipulation software, e.g. direct change of brightness or blurring for shadow boundaries, requires a considerable amount of manual adjustment to align the appearances of modified shadow to the original shadow. This alteration requires delicate editing steps and unavoidably results in unnatural artefacts around shadow boundaries. These features of shadow editing are not available in current image manipulation software, e.g. Photoshop and GIMP.

Recent shadow removal work (Shor and Lischinski, 2008; Su and Chen, 2010; Wu et al., 2007) presents some basic examples of shadow editing including complete removal, duplication, distortion and sharpness adjustment of the original shadow. The applications of this work only apply simple image manipulation to the original shadow matte and do not provide a manipulatable model for arbitrary shadow modification. A shadow editing system is first proposed in (Mohan et al., 2007) which requires users to

specify some boundary points of shadow (e.g. Fig. 2-13(a)). Their shadow edge model is manipulatable and supports controls for sharpness, darkness and shape of shadows. However, users are only allowed to move the specified boundary points of shadow which limits its range of amendable shapes. Users are also not allowed to add new shadow segments. In summary, natural shadow editing for arbitrary shapes and properties is still not explored.

2.10 Summary

Approaches to shadow removal can be categorised as either automatic (Finlayson et al., 2007, 2009, 2002a,b, 2006; Guo et al., 2012; Yang et al., 2012) or user-aided (Arbel and Hel-Or, 2011; Liu and Gleicher, 2008; Shor and Lischinski, 2008; Su and Chen, 2010; Wu et al., 2007). The differentiation between fully automatic or user-aided relates to initial detection of the shadow – with removal itself (after detection) being a largely automatic task. In any case, both removal and detection are ill-posed problems and difficult to reliably achieve.

As for automatic shadow removal, intrinsic image based methods are one such popular approach to the problem, such as (Finlayson et al., 2007, 2009, 2002a,b, 2006; Yang et al., 2012). The decomposition of intrinsic images provides reflectance and/or shading information but can be unreliable leading to over-processed results. They generally assume that the illumination change leads to smooth intensity change and the colours of the scene can be described by a few primitive colours. Other approaches (Guo et al., 2012; Huang and Chen, 2009; Lalonde et al., 2010; Zhu et al., 2010) detect shadows by learning shadow features. However, detection is constrained by the range of training data and quality of image edge detection and segmentation there-in. (Huang and Chen, 2009; Lalonde et al., 2010; Zhu et al., 2010) detect shadows by classifying edges in an image according to shadow features such as changes in intensity, texture and colour ratio. (Guo et al., 2012) adopts similar features but detects shadows by classifying segments in an image and pairing shadow and lit segments globally. (Guo et al., 2012) shows more robust shadow detection due to its use of segment pairing. Some methods require additional controllable light sources to capture shadowless objects, e.g., by capturing images with a light source placed in different positions (J. J. Yoon and T.J.Ellis, 2002) or by comparing flash and no-flash image pairs (Drew et al., 2006). However, active lighting restricts the applicable type of scenes - as moving lights around and using special lighting setups outdoors is often not practical. Some methods adopt optical filters to obtain multi-spectral images for illumination detection, e.g. by comparing NIR and RGB images (Salamati et al., 2011) and by comparing RGB and single-colour-filtered image (Finlayson et al., 2007), but these methods assumes some special scenarios, e.g. sunlight and non-black surfaces. They are thus not applicable to the removal of normal

single images.

User-aided methods generally achieve higher accuracy in shadow detection at the practical expense of varying degrees of manual assistance. Wu et al. require extensive user input where the user needs to define multiple regions of shadow, lit area, uncertainty and exclusion. They apply a Bayesian optimisation to derive a shadow matte and a shadow-free image. Others (Liu and Gleicher, 2008; Su and Chen, 2010) require fine input defining the shadow boundary. Liu and Gleicher propose a curve fitting method and a global alignment of gradients to acquire shadow scales but have issues when relighting the umbra and can introduce artefacts at uneven boundaries. Shor and Lischinski detect shadow using image matting from a grown shadow seed. They only require one shadow pixel as input, but have limitations in cases where the other shadowed surfaces are not surrounded by the initially detected surface or when the penumbra is too wide. Arbel and Hel-Or apply a model of a thin-plate fitting to fit the intensity surface. They require users to specify multiple texture anchor points to detect a shadow mask but this can increase when shadow regions are multiple and scattered. Su and Chen estimate shadow scales by using gradient priors based optimisation. Their gradient alignment for intensity samples allows for rougher user input compared with (Liu and Gleicher, 2008; Wu et al., 2007). In addition, they provide a healing tool for users to manually amend artefacts on highly-curved shadow boundary segments but can often fail when the surface texture is strong or irregular. Shor and Lischinski apply native in-painting for penumbra recovery which result in penumbra artefacts. Most others (Arbel and Hel-Or, 2011; Liu and Gleicher, 2008) assume highly-constrained curve or surface functions for illumination change which limit their range of removable shadows. To summarise, the features and requirements of the reviewed shadow removal methods are compared in Tab. 2.1.

To date, most shadow removal methods (Arbel and Hel-Or, 2011; Su and Chen, 2010; Wu et al., 2007) have been evaluated by visual inspection on some selected images – with only a few exceptions performing quantitative evaluation (Guo et al., 2012; Shor and Lischinski, 2008). This is in part due to a lack of high-quality, varied, and public ground truth data. The difficulty of collecting such a data set is highlighted in the first public data set (Guo et al., 2012), with the appearance of some global illumination changes and mis-registration between the shadow and shadow-free images being a difficult factor to control. This makes quantitative testing on such data unreliable. Another desirable property as yet not explored is the categorisation of shadows. The attributes of shadows are important to consider because these different shadow types can present their own unique challenges. The categorisation also allows us to evaluate different algorithm performance in a range of scenarios and scene types.

Natural shadow editing for arbitrary shapes and properties is still unexplored. The demanding features include the editing controls for colour, softness, darkness of shadows

Method	User Input	Illumination Preserving	Texture Preserving	Colour Correction	Special Setup
Finlayson et al. (2002a)	None	No	No	No	None
Finlayson et al. (2006)	None	No	No	No	None
Finlayson et al. (2009)	None	No	No	No	None
Fredembach and Finlayson (2005)	None	No	No	No	None
Fredembach and Finlayson (2006)	None	Yes	No	No	None
Yang et al. (2012)	None	No	Yes	No	None
Guo et al. (2012)	None	Yes	Yes	No	None
J. J. Yoon and T.J.Ellis (2002)	None	No	Yes	No	Flash light
Drew et al. (2006)	None	No	Yes	No	Flash light
Salamati et al. (2011)	None	Yes	Yes	No	NIR/RGB Camera
Mohan et al. (2007)	multiple shadow boundary points	No	Yes	No	None
Wu et al. (2007)	multiple strokes for each shadow segment	Yes	Yes	No	None
Liu and Gleicher (2008)	precise boundary mark for each shadow segment	Yes	Yes	Yes	None
Shor and Lischinski (2008)	single click for each shadow segment	Yes	No	No	None
Su and Chen (2010)	rough boundary mark for each shadow segment	No	Yes	No	None
Arbel and Hel-Or (2011)	multiple clicks for each shadow segment	Yes	Yes	Yes	None

Table 2.1: *Feature comparison of recent shadow detection and removal methods (methods for performing shadow detection* only *– and not removal – are not listed in this table): "Illumination Preserving" refers to the ability to preserve the original illumination in the lit area. "Texture Preserving" refers to the preservation of the correct surface texture under the penumbra after removal. "Colour Correction" refers to the ability to correct colour artefacts caused by image post-processing after removal. "Special Setup" is the requirement for special devices other than a trivial camera.*

as well as their shapes. Although there is an existing method for editing the original shadow region by using some sparse control points (Mohan et al., 2007), more flexible approaches for re-defining new or existing shadow shapes are still not available.

Chapter 3

Texture-Preserving Shadow Removal

Overview

This chapter presents a novel user-aided method for texture-preserving shadow removal from single images which only requires simple user input. Compared with the state-of-the-art methods, this algorithm addresses limitations in uneven shadow boundary processing and umbra recovery. We first detect an initial shadow boundary by growing a user specified shadow outline on an illumination-sensitive image. Interval-variable intensity sampling is introduced to avoid artefacts raised from uneven boundaries. We extract the initial scale field by applying local group intensity spline fittings around the shadow boundary. Bad intensity samples are replaced by their nearest alternatives based on a log-normal probability distribution of fitting errors. Finally, we use a gradual colour transfer to correct post-processing artefacts such as gamma correction and lossy compression. Compared with state-of-the-art methods, we offer highly user-friendly interaction, produce improved umbra recovery and improved processing given uneven shadow boundaries.

3.1 Introduction

This chapter focuses on user-aided single image shadow removal. User-aided methods generally achieve better shadow detection and removal at the cost of user input. Many of them (Arbel and Hel-Or, 2011; Liu and Gleicher, 2008; Mohan et al., 2007; Wu et al., 2007) are also texture-preserving. Although these past works preserve penumbra texture, umbra removal and uneven boundary processing are still problematic. The umbra is the darkest part of the shadow while the penumbra is the wide outer boundary with a non-linear intensity change between the umbra and lit area. The penumbra scale

is non-uniform and within the penumbra shadowed surface textures generally become weaker. Assuming a uniform umbra, (Liu and Gleicher, 2008) often introduces over-saturation artefacts in some cases. Uneven shadow boundaries can affect penumbra detection and scale estimation. Contrast artefacts can also appear in shadow areas due to image post-processing (Arbel and Hel-Or, 2007, 2011; Liu and Gleicher, 2008). Most user-aided methods to assist boundary detection (Arbel and Hel-Or, 2007; Liu and Gleicher, 2008; Wu and Tang, 2005) require complex user input for locating the shadow boundary. We propose a method that requires one rough stroke to mark an umbra sample. It can process uneven shadow boundaries and achieves better umbra removal compared with past work.

In this chapter, the shadow effect is represented as an additive scale field in the log domain as described in Eq.2.7.

The major contributions in this chapter are as follows:

1) Easy user input: Previous work, e.g. (Arbel and Hel-Or, 2007, 2011; Liu and Gleicher, 2008; Wu and Tang, 2005; Wu et al., 2007), requires precise user-inputs defining the shadow boundary. Our method only requires an umbra segment highlighted by one rough stroke and grows this on an illumination-sensitive image to obtain initial shadow boundaries.

2) Interval-variable sampling: Previous work, e.g. (Arbel and Hel-Or, 2007, 2011; Liu and Gleicher, 2008), applies interval-fixed sampling around the penumbra that causes artefacts near uneven shadow boundaries. To address this, we develop an interval-variable sampling according to shadow boundary curvature.

3) Local group optimisation for picked samples: Inspired by Arbel's (Arbel and Hel-Or, 2007, 2011) and Liu's (Liu and Gleicher, 2008) spline-fitting approach, we propose a local group optimisation that balances curve fitness and local group similarity. Unlike past work, we filter inferior samples that are replaced with their closest alternatives according to a log-normal probability distribution. This reduces shadow removal artefacts.

4) Gradual colour transfer: Post-processing effects cause inconsistent shadow removal compared with the lit areas both in tone and in contrast. We make use of statistics in thin penumbra boundaries and the shadow scale field to correct these issues.

3.2 Proposed Method

Figure 3-1: *Shadow removal overview. (a): user input indicated by a purple curve, (b): initial shadow boundary detection indicated by the purple mask (§3.2.1), (c): interval-variable sampling lines indicated by the white lines (§3.2.2), (d): shadow removal by estimating the scale field of the shadow area (§3.2.2). Note that the area's contrast still appears different from its surroundings. (e): image correction by our gradual colour transfer (§3.2.3). The input image is taken from (Arbel and Hel-Or, 2007).*

Given an input image and a user specified umbra segment (Fig. 3-1(a)), we detect the initial shadow boundaries by expanding the given umbra segment on an illumination image using an active contour (Fig. 3-1(b), §3.2.1). To keep boundary details, we sample pixel intensities along the sampling lines perpendicular to the shadow boundary (Fig. 3-1(c), §3.2.2) at variable intervals. We perform a local group optimisation to estimate the illumination change which refines shadow boundary detection and provides an initial scale field. According to an adaptive sample quality threshold, sampling lines with bad samples are replaced by their nearest neighbours and a later local group optimisation is applied to them. Finally, we relight the shadow area using our scale field (Fig. 3-1(d), §3.2.2) and correct post-processing artefacts using our gradual colour transfer (Fig. 3-1(e), §3.2.3).

3.2.1 Initial Shadow Boundary Detection

Determining the initial shadow boundary is the first step of penumbra detection and is required in many previous methods including (Arbel and Hel-Or, 2007, 2011; Liu and Gleicher, 2008). In this subsection, we explain how to derive an initial shadow boundary from a given rough umbra sample segment.

(a) src (b) input (c) $\mathcal{C}_1$ (d) $\mathcal{C}_2$ (e) $\mathcal{C}_3$ (f) $\mathcal{C}_4$ (g) $\mathcal{F}$

Figure 3-2: *An illumination-sensitive image $\mathcal{F}$ is fused from the four colour channels of the image (a) by using the intensity statistics of the given umbra region (b) inside the purple boundary. We grow the region by using the image $\mathcal{F}$ to the entire shadow area as indicated in (g) by the purple mask. Sub-figures (c-f) are the region growing results on the corresponding four single channels. The input image is taken from (Liu and Gleicher, 2008).*

Inspired by (Katramados et al., 2009), we fuse four normalised candidate illumination-sensitive channels from different colour spaces into an illumination image. The chosen channels are: the V channel ($\mathcal{C}_1$) of HSV space, the V channel ($\mathcal{C}_4$) of LUV space, and the Y channel ($\mathcal{C}_2$) and Cb channel ($\mathcal{C}_3$) of YCbCr space. We measure the confidence values of each candidate channel using an exponential incentive function φ representing the textureness of each of their umbra sample segments:

$$\varphi(x) = x^{-\lambda}(\lambda > 0) \tag{3.1}$$

where x is the pixel intensity, λ (default value 5) determines the steepness of the incentive function. Lower textureness is preferred as it means higher intensity uniformity of umbra segment. The textureness is measured by standard deviation of intensities. The fused image $\mathcal{F}$ is computed as a weighted sum of each normalised candidate channel $\mathcal{C}_l$ as follows:

$$\mathcal{F} = (\sum_{l=1}^{4} \mathcal{C}_l \varphi(\sigma_l))/(\sum_{l=1}^{4} \varphi(\sigma_l)) \tag{3.2}$$

where l is the channel index, σ_l is the standard deviation of the umbra sample intensities of $\mathcal{C}_l$. To avoid texture noise, we apply a bilateral filter (Paris and Durand, 2009) to $\mathcal{F}$ first. We grow a sparse-field active contour (Whitaker, 1998) on the fused image to detect the initial shadow boundary. As shown in Fig. 3-2, region-growing using the fused image is more robust than relying on one channel.

3.2.2 Scale Field Estimation

This subsection describes our scale field estimation from initial shadow detection. Shadow effects are represented by varying (or different scale) intensity values. Using a scale field better represents the penumbra and umbra variations, and is used to relight the shadow area using Eq. 2.7. In § 3.2.2, we first sample the log domain pixel intensity along the sampling lines perpendicular to the shadow boundary. In § 3.2.2, we adopt a local group spline fitting optimisation through the measured sampling line pixel intensities to estimate sparse scales from the initial intensity samples. We replace bad intensity samples with their nearest alternatives and re-optimise for them. We spread the sparse scales to dense scales followed by a gradual colour transfer in § 3.2.3 that adjusts the colour and texture of initial shadow removal.

Interval-Variable Sampling

According to Eq. 2.7, the logarithms of the original image are supplied for sampling. We sample pixel intensities along the lines perpendicular to the initial shadow boundary as shown in Fig. 3-4(c). Uneven boundaries can result in non-smooth vector normal estimations along the shadow boundary. To overcome this, we apply cubic spline smoothing

(a) original (b) over-sampling (c) suitable sampling

Figure 3-3: *(a) is the original patch. (b) is the result of sampling every shadow boundary pixel, which results in shadow removal noise near the boundary. (c) is the noise-free result with a larger sampling interval.*

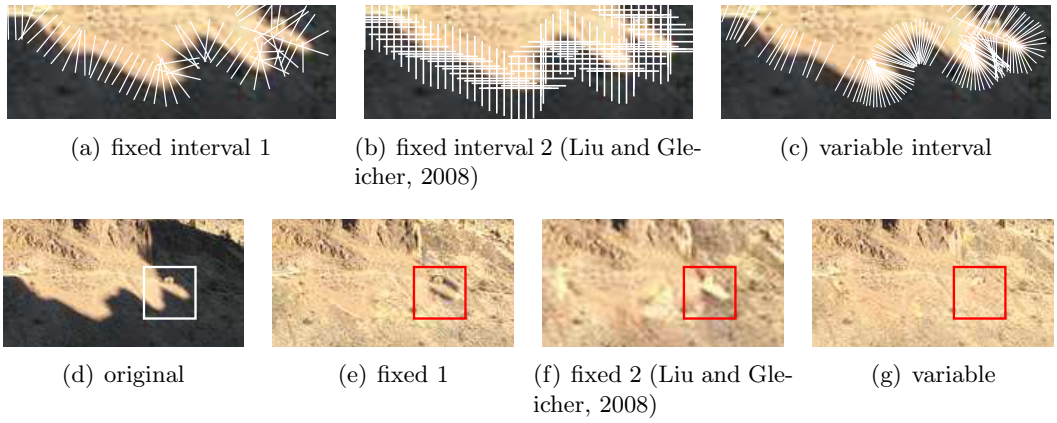
to the initial boundary points before we compute their normals and curvatures. Under-sampling along the boundary neglects sharp details and causes artefacts as shown in Fig. 3-4 while over-sampling incurs penumbra removal noise due to texture as shown in Fig. 3-3. More sparse pixel scales are computed for curvy boundary parts for precise in-painting. To avoid texture artefacts, we apply a bilateral filter (Paris and Durand, 2009) to the input image before sampling. Unlike past work (Arbel and Hel-Or, 2007, 2011; Liu and Gleicher, 2008; Mohan et al., 2007), we do not adopt a fixed sampling line interval, e.g., one sampling line per boundary pixel. Our method adjusts the sampling interval according to the curvature of the smoothed boundary. The interval is the same for all RGB channels. We set a curvature accumulator for shadow boundary points and accumulate along the boundary. We place sampling marks and reset the accumulator when the curvature sum reaches a threshold ξ (default value 0.05). We achieve this by adopting Eq. 3.3. We limit the absolute curvature of each boundary point with an upper limit ξ and compute a cumulative sum array Q of the saturated absolute curvature array. To determine the sampling interval, we choose boundary points where the sampling lines pass through as follows:

$$\begin{cases} \tilde{Q}_m = \lfloor Q_m/\xi \rfloor & (m \leqslant N, m \in \mathbb{N}) \\ D_n = \tilde{Q}_{n+1} - \tilde{Q}_n & (n \leqslant N-1, n \in \mathbb{N}) \end{cases} \tag{3.3}$$

where N is the number of boundary points, m and n specify the index of boundary points, $\tilde{Q}$ is the array of the quantized and normalised cumulative sum Q, D is the adjacent element differences array of $\tilde{Q}$. To get the interval of sampling marks, we set the marks for the first and last boundary points and all the points in D with non-zero values. If the boundary is a straight line, the initial interval is fixed up to a maxima (five boundary points). As shown in Fig. 3-4, our variable sampling interval avoids penumbra removal artefacts around sharp boundary parts.

Illumination Variance Estimation

Having obtained sparse intensity samples at different positions along the boundary, our goal is to find illumination scaling values inside the umbra, penumbra and lit area. We model the illumination scale change S_i for each i^{th} intensity sample of each RGB



(a) fixed interval 1 (b) fixed interval 2 (Liu and Gleicher, 2008) (c) variable interval

(d) original (e) fixed 1 (f) fixed 2 (Liu and Gleicher, 2008) (g) variable

Figure 3-4: *The white lines in (a), (b), and (c) are the sampling lines of the fixed interval using boundary-perpendicular sampling, the fixed interval using horizontal/vertical sampling in (Liu and Gleicher, 2008), and our boundary-perpendicular variable interval sampling respectively. (d) is the original image. (e), (f), and (g) are the corresponding shadow removal results of the three sampling methods respectively.*

channel as follows (see also Fig. 3-5):

$$S_i(x) = \begin{cases} K & x \leq x_1 \\ f(x) & x_1 < x \leq x_2 \\ 0 & x > x_2 \end{cases} \tag{3.4}$$

where x is a pixel location within the sampling line, x_1 and x_2 determine the start and

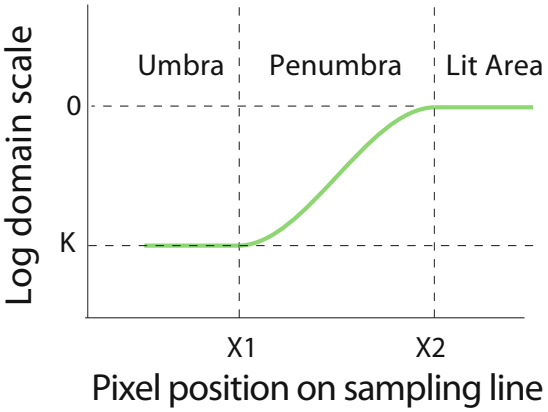


Figure 3-5: *Shadow scale model*

end of the penumbra area respectively, and K is a negative scale constant for sample points within the umbra area ($x < x_1$). The constant 0 is assumed for the lit area piece ($x > x_2$) as this falls inside a lit area of the image and does not require re-scaling. In order to solve x_1, x_2, K, and f, we solve a piece-wise function G_i parametrised by v_1,

v_2, v_3, and v_4 (recall that our intensity samples are in the log domain):

$$\begin{cases} G_i(x) = S_i(x) + v_4 \\ f(x) = v_1 B(v_2(x - v_3)) \\ B(y) = -0.25y^3 + 0.75y - 0.5 \\ \begin{bmatrix} x_1 & x_2 \end{bmatrix} = v_3 + v_2^{-1} \begin{bmatrix} -1 & 1 \end{bmatrix} \\ K = -v_1 \end{cases} \tag{3.5}$$

where B is a cubic shape function (a sinusoidal function here also produces adequate results) and y is the input. Our solution is thus reduced to solving v_1, v_2, v_3, and v_4 for each RGB channel (twelve in total). Illumination of each channel may vary differently while penumbra boundaries of three channels are usually the same. We thus assume a common penumbra width and position for each channel, determined by v_2 and v_3 respectively. We formulate solving the eight remaining parameters as an optimisation problem which balances curve fitness and local group fitting similarity. We minimise the energy function E_i for the i$^{\text{th}}$ sampling line as follows:

$$E_i = \alpha_1 E_{fit}(V_i, Z_i) + \alpha_2 E_{gs}(V_i, V_{i-1}, \ldots, V_{i-r-1}) \tag{3.6}$$

where r is the number of members in a local sampling line group (default value 5), α_1 and α_2 are two balancing weights (default values 1 and 0.2 respectively), E_{fit} measures the sum-of-square fitting error between the three piece-wise functions G_i (defined by parameter vector V_i) and the original three-channel intensity sample matrix Z_i, E_{gs} measures the parameter similarity between the neighbouring members of a local group. In practice, penumbra width affects the removal quality most significantly. We thus only compare the similarity of the shared v_2 in Eq. 3.5. E_{gs} is defined as follows:

$$\begin{cases} V_i = [v_{R1}, v_{G1}, v_{B1}, v_{RGB2}, v_{RGB3}, v_{R4}, v_{G4}, v_{B4}] \\ E_{gs}(V_i, V_{i-1}, \ldots, V_{i-r-1}) = \left(\frac{1}{b_i} - \frac{\sum_{k=1}^{r} \varphi(\dot{e}_{i-k})}{\sum_{j=1}^{r} \varphi(\dot{e}_{i-j}) b_{i-j}} \right)^2 \end{cases} \tag{3.7}$$

where v_{R1}, v_{G1} and v_{B1} are the v_1 for each channel, v_{R4}, v_{G4} and v_{B4} are the v_4 for each channel, v_{RGB2} and v_{RGB3} are the shared v_2 and v_3 for all channels, b_i is the v_{RGB2} of the parameter vector V_i, $\dot{e}_u$ indicates the fitting error of the previous u$^{\text{th}}$ fittings, φ is the function defined in Eq. 3.1. We solve this using a sequential quadratic programming algorithm (Nocedal and Wright, 2006).

However, interval-variable sampling can not always guarantee good sample quality. Strong surface textures introduce more significant intensity change than illumination change. Unlike past work(Arbel and Hel-Or, 2007, 2011; Liu and Gleicher, 2008; Mohan

(a) original (b) removal (c) correction

Figure 3-6: *Gradual colour transfer. (b) is the initial shadow removal of (a) and is corrected by our gradual colour transfer as shown in (c).*

et al., 2007), we ignore sampling lines with higher fitting errors and pick their most suitable neighbours. Based on our empirical tests on various images, we model the initial fitting error distribution as a log-normal probability distribution. As the initial fitting error is distributed log-normally, we can convert it to its corresponding normal distribution by taking the exponential of it. According to the empirical 3-sigma rule of normal distribution, we ignore sampling lines with fitting errors higher than a threshold $\mu+\sigma$ which accounts for 15.8% of all samples where μ and σ are the mean and standard deviation of errors. For each sampling line with bad samples, we only attempt to pick its nearest sampling lines within a short distance, i.e. no further than its neighbouring chosen sampling lines. To evaluate the replacements' quality and compute their sparse scales, we apply the same optimisation method described previously.

After optimisation, we obtain sparsely distributed scales, defined in Eq. 2.7, of all sampled pixels inside and around the penumbra area. We interpolate the scales within the penumbra area using linear interpolation and extrapolate the other scales in the lit and umbra area using in-painting (Bertalmio et al., 2000).

3.2.3 Gradual Colour Transfer

In practice, the theoretical shadow effect formulation does not often hold. Image acquisition devices usually apply post-processing, e.g. Gamma correction. Lossy compressions, e.g. JPEG, are also common in images such that compression artefacts (e.g. affecting contrast) in the shadow area become noticeable when removal is applied. To address this, we extend the colour transfer in (Reinhard et al., 2001) with a scale field S_m. We compute the normalised scale increase h_i of the i$^{\text{th}}$ sampling line according to Eq. 3.4 as follows:

$$h_i(x) = \frac{exp(f_i(x)) - exp(K_i)}{1 - exp(K_i)} \tag{3.8}$$

where x is the location of pixel of a sampling line, K_i and f_i are respectively the lit constant K and the cubic function piece of the i$^{\text{th}}$ sampling line. We apply the same method described in § 3.2.2 to interpolate and extrapolate sparse scale increase values computed by Eq. 3.8 to form a dense scale increase field S_m. We convert the initial shadow removal image from RGB space to LAB space. For each LAB channel, we

compute the mean μ_u and deviation σ_u of the umbra side pixel intensities near the penumbra as the source and we compute the mean μ_t and deviation σ_t of the lit side pixels near the penumbra as the target. We adjust the initial removal image channel L to the final image channel $\tilde{L}$ as follows:

$$\begin{cases} \mu_s(x,y) = \mu_u + S_m(x,y)(\mu_t - \mu_u) \\ \sigma_s(x,y) = \sigma_u + S_m(x,y)(\sigma_t - \sigma_u) \\ \tilde{L}(x,y) = \mu_t + (L(x,y) - \mu_s(x,y))\sigma_t/\sigma_s(x,y) \end{cases} \tag{3.9}$$

where x and y are the coordinates of pixels, μ_s and σ_s are the fields of gradual source mean and deviation. We show an example of colour transfer in Fig. 3-6.

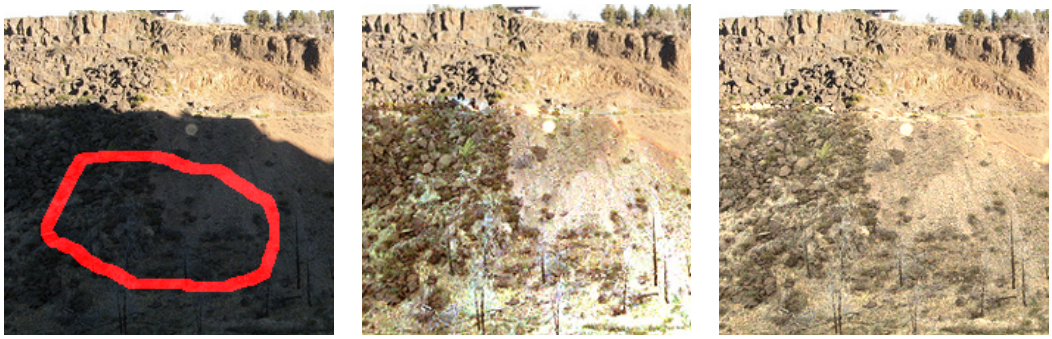
3.3 Evaluation

In our experiments we visually compare our shadow removal results with state-of-art methods using the same input images and some other representative images. Our method has been tested on various photos from existing work and additional representative photos. Our method is highly user-friendly as shown in Fig. 3-1 and produces state-of-art quality shadow removal as shown in Fig. 3-1, 3-3, 3-4, 3-7, 3-6, 3-8, and 3-9. For each shadow area, we require one stroke, shown as red curves in figures, marking an characteristic umbra segment. The stroke can be very rough and does not necessarily need to follow the shadow boundary. We also overcome some cases that past work fail to handle.

In our comparisons, we focus primarily on more recently studied texture-preserving methods as opposed to older texture-lossy ones such as in-painting (Finlayson et al., 2006; Fredembach and Finlayson, 2006; Shor and Lischinski, 2008). Our previous figure (Fig. 3-4) has already highlighted the issues caused by uneven boundary processing. This compares our variable sampling interval – which is boundary-perpendicular – against our own test of fixed interval boundary-perpendicular sampling and the fixed interval vertical and horizontal sampling method in (Liu and Gleicher, 2008).

Compared with (Liu and Gleicher, 2008), our method reduces the red light component that passes through the semi-transparent leaf as illustrated in Fig. 3-7. The redness of different parts of the leaf is different and thus the amount of passed red light is not uniform. To handle this, we assume non-uniform scales in the umbra, i.e. different umbra scale constants as mentioned in § 3.2.2, and process the RGB channels separately based on a shared penumbra area. Our interval-variable sampling (see Fig. 3-7(c)) removes some minor residual shadow fragments near the shadow boundary (noticeable in Fig. 3-7(b)). In Fig. 3-7(d), the shadowed sandy surface is consistently recovered. As the surface in the lit area is not saturated, the same surface beneath

(a) original from (Liu and Gleicher, 2008) (b) result from (Liu and Gleicher, 2008) (c) our result



(d) original from (Liu and Gleicher, 2008) (e) result from (Liu and Gleicher, 2008) (f) our result

(g) original from (Mohan et al., 2007) (h) result from (Mohan et al., 2007) (i) our result

Figure 3-7: *From (a), the red light component that passes through semi-transparent object is still in (b) and is reduced by our method in (c). Ours also removes shadow residuals near the highly uneven boundary in (b). The shadow removal result of (d) shown in (e) results in oversaturation of the umbra. Also, the texture and colour variations across the shadow boundary are not smooth and consistent. We overcome this issue as shown in (f). Our removal result of (g), as shown in (i), is improved over (Mohan et al., 2007) as shown in (h).*

the shadow should also not be saturated in Fig. 3-7(e). Fig. 3-7(f) shows our method avoids over-saturation artefacts and achieves more consistent texture and smoother colour variance across the shadow boundary. In Fig. 3-7(h), the result from (Mohan et al., 2007) appears darker in both the shadow and lit areas. Our result shows a consistently coloured texture between the lit area and shadowed area. Figs. 3-8 and 3-9 demonstrate our results on images with various textures, reflectances, and shadow boundaries. Fig. 3-8(a) shows the removal of soft shadow cast on the curved surface and the texture consistency. Fig. 3-8(b) reveals our smooth and texture-consistent shadow removal applied to the earlier example illustrated in Fig. 3-4. The colour of the tree bottom of the hill is consistent with the trees at the hill top. The smoothness of the hill-side's colour and texture are also recovered. In Fig. 3-8(d), Fig. 3-8(e), Fig. 3-8(c), and Fig. 3-8(f), the texture and self-shadows are kept after removal.

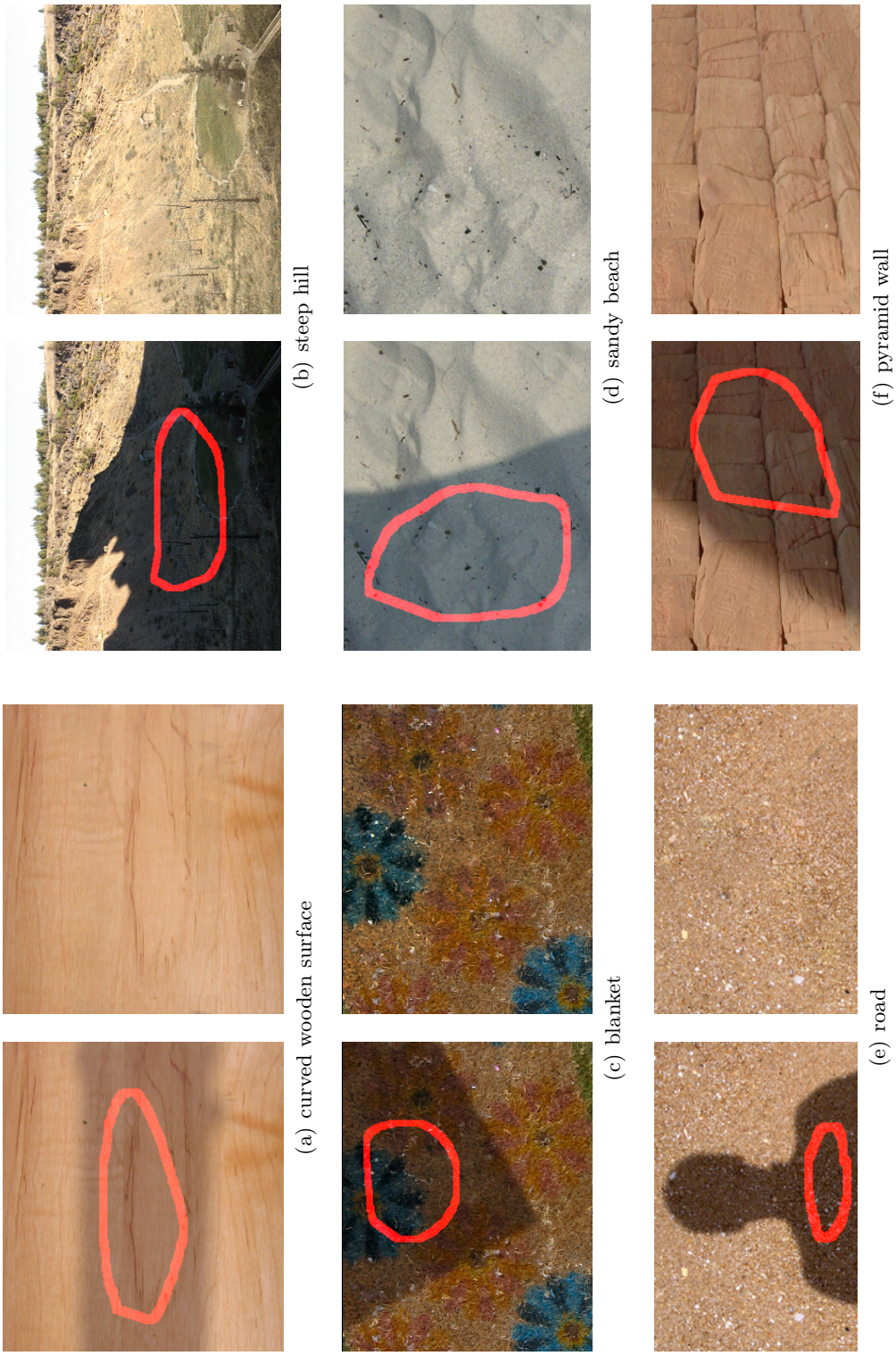


Figure 3-8: *Demonstrations. Left: original images; Right: our results. The original images in (a) and (c) are from (Arbel and Hel-Or, 2007) and the original images in (b) and (f) are from (Liu and Gleicher, 2008) and (Mohan et al., 2007) respectively.*

However, our method still shows some limitations in highly complex cases (still unsolved in state-of-the-art work). In Fig. 3-9(a), the river bed is recovered but the ripple highlights in the shadow area are missing. These highlights are mainly produced by light reflection on the wavy water surface and the light refraction from the river bed. As the directed light is blocked in the shadow area, these complex affects cannot be recovered by simple relighting. In a challenging case in Fig. 3-9(b), the piecewise penumbra illumination change model does not fit well with the actual penumbra intensity change resulting in some band-like artefacts in the penumbra. In Fig. 3-9(c), the global colour correction in this chapter can cause artefacts when the colour distribution of lit and shadow sides are dissimilar.

(a) rocky river beach

(b) penumbra band-like artefact

(c) global colour correction artefact

Figure 3-9: *Failure cases. Left: original images; Right: our results.*

3.4 Summary

We have presented a user-friendly texture-preserving shadow removal method that overcomes some common limitations from the past work. Specifically, our method retains shadowed texture and preforms well on highly-uneven shadow boundaries, non-uniform umbra illumination, and non-white lighting. Our main technical contributions are: (1) highly user-friendly input; (2) interval-variable sampling; (3) local group optimisation; and (4) gradual colour transfer. This method was one of the best in 2013, however it can not handle complex shadow scenes well and users have to supply many strokes for complex scenes with multiple shadow regions. In the following Chapter 4, an artefact-resistant shadow removal method is proposed to remove the shadows in more complex scenes and minimise the artefacts caused by piece-wise illumination change model and global colour correction. These complex shadows include soft shadows, and shadows cast on strong texture backgrounds. In addition, a quantitative shadow removal evaluation is performed.

Publications

The following publication is related to this chapter:
Gong, H., Cosker, D., Li, C., and Brown, M. (2013). User-aided single image shadow removal. In *IEEE Proc. International Conference on Multimedia and Expo (ICME)*

Chapter 4

Artefact-Resistant Shadow Removal

Overview

This chapter presents a novel user-aided method for texture-preserving shadow removal from single images requiring only simple user input. Compared with the state-of-the-art approaches, this algorithm offers the most flexible user interaction to date and produces more accurate and robust shadow removal under thorough quantitative evaluation. Compared with our texture-preserving shadow removal presented in Chapter 3, this algorithm provides an easier user interaction for scenes with multiple shadow regions and significantly improves the shadow removal results of difficult scenes such as soft shadows and shadows cast on strong texture backgrounds. Shadow masks are first detected by analysing user specified shadow feature strokes. Sample intensity profiles with variable interval and length around the shadow boundary are detected next, which avoids artefacts raised from uneven boundaries. Texture noise in samples is then removed by applying local group bilateral filtering, and initial sparse shadow scales are estimated by fitting a piecewise curve to intensity samples. The remaining errors in estimated sparse scales are removed by local group smoothing. To relight the image, a dense scale field is produced by in-painting the sparse scales. Finally, a gradual colour correction is applied to remove artefacts due to image post-processing. Using a public evaluation data set, we quantitatively and qualitatively demonstrate our method to outperform current leading shadow removal methods.

4.1 Introduction

In this chapter, a semi-automatic method is proposed for high-quality shadow removal using user-defined flexible single strokes covering the shadow and lit pixels. Compared

with the user interaction adopted in Chapter 3, the user interaction used in this chapter does not require users to mark every shadow region for complex scenes. Instead, users are only required to supply fewer strokes indicating some samples of related shadow and lit pixels. Given detection, our method produces accurate shadow removal optimised for robust penumbra recovery. Using a public shadow removal ground truth data set (Guo et al., 2012), our solution is quantitatively evaluated against our texture-preserving shadow removal (Chapter 3) and the other leading methods and demonstrates notably improved performance. Numerous visual comparisons of our method versus existing methods are also presented, demonstrating qualitatively more pleasing results. Compared with our shadow removal solution in Chapter 3, we have made the following improvements: *a*) an intelligent sample selection and de-noising; *b*) a more efficient scale estimation and a new boundary artefact remover; *c*) an improved colour correction for complex background of surface; *d*) a simplified user interaction for complex scenes; *e*) an improved quantitative evaluation based on an existing shadow removal data set (Guo et al., 2012) . Therefore, our approach represents a state-of-the-art technique for shadow removal with a thorough evaluation against the current leading approaches.

As mentioned previously in the background chapter (Chapter 2), most shadow removal methods (e.g. (Arbel and Hel-Or, 2011; Liu and Gleicher, 2008; Su and Chen, 2010; Wu et al., 2007)) have only been evaluated by visual inspection on some selected images. The only public ground truth data set for shadow removal (Guo et al., 2012) is not reliable enough. The data set is small in size and some of their shadow-free ground truth shows inconsistent illumination compared with the lit area of their corresponding shadow images. Also, their error measurement is variant to the size and darkness of shadows.

In this chapter, the shadow effects is represented as a Hadamard product of a shadow scale layer and a shadow-free image as previously described in Eq. 2.4.

4.1.1 Contributions

Given our overview of state-of-the-art approaches, three main contributions are proposed:

1) Easy user input: Previous work, e.g. (Arbel and Hel-Or, 2011; Liu and Gleicher, 2008; Su and Chen, 2010; Wu et al., 2007), requires precise user-input defining the shadow boundary. Our user interaction in Chapter 3 require users to mark every shadow region which can be time-consuming for images contain multiple shadows. Our method only requires users to define a single simple rough strokes covering related shadow and lit pixels – without the need to differentiate between samples in shadow and lit areas. Users are also not required to mark every shadow region.

2) Intelligent sampling: Adaptive sampling with variable intervals and lengths is

proposed to address shadow boundary artefacts in past work (Arbel and Hel-Or, 2011; Liu and Gleicher, 2008), which uses fixed intervals and lengths. Unlike past work (Arbel and Hel-Or, 2011; Liu and Gleicher, 2008; Su and Chen, 2010), unqualified samples are intelligently filtered. These can affect the quality of shadow scale estimation, e.g. samples with high noise or sampling lines passing through boundaries caused by occlusions or strong background texture. The solution in Chapter 3 removes the unqualified samples according to their curve fitting errors. However, this requires the redundant curve fitting to be done before it can evaluate the quality of the sample. The sample selection in this chapter evaluates the sample quality before fitting which is obviously faster.

3) Robust scale estimation Fast local group processing is proposed for selected samples and initially estimated scales to improve smoothness of shadow removal. Post-processing effects cause inconsistency in shadow corrected areas compared with the lit areas both in tone and contrast. Without introducing chromatic artefacts, e.g. the colour correction in Chapter 3, colour-safe correction is proposed to amend the scales.

To summarise, this chapter presents several solutions to improve shadow removal quality, and these have been quantitatively verified using robust error measurement against the leading state-of-the-art data set in this area (Guo et al., 2012). In addition, we also consider performance on different *categories* of shadow, i.e. the type of shadow being cast.

4.2 Artefact-Resistant Shadow Removal

In this section, our algorithm is first described in brief before being expanded on with technical details for each of its components. Our algorithm consists of four steps (see Fig. 4-1):

1) Pre-processing (§4.2.1) A shadow mask is detected (Fig. 4-1(b)) using a KNN classifier trained from K-Means clustered data from user inputs (e.g. Fig. 4-1(a)). A *fusion image* is generated, which provides an illumination-insensitive layer, by fusing the channels of YCrCb colour space and de-noising (Fig. 4-1(c)).

2) Intensity sampling (§4.2.2) Intensity profiles are obtained for sampling lines perpendicular to shadow boundaries. *Poor* samples are filtered based on similarity of illumination change (Fig. 4-1(d)) and de-noised using directional bilateral filtering (Fig. 4-1(e)).

3) Estimation of shadow scale and relighting (§4.2.3) Given the filtered intensity samples, these are fit through and relit (Fig.4-1(f)) using a piece-wise cubic curve and a boundary image of the samples (Fig. 4-1(e)). Any remaining boundary artefacts are removed using directional scale suppression (Fig. 4-1(g)) over the boundary image. Fitted sparse scales are propagated (Fig. 4-1(h-i)) to generate a dense scale field (Fig. 4-

1(j)). Shadows are then removed (Fig. 4-1(k)) by inverse scaling using this dense scale field.

4) Gradual colour correction (§4.2.4) Any remaining shadow removal artefacts due to image post-processing are finally treated with our colour correction (Fig. 4-1(l)). This uses statistics around penumbra boundaries and the shadow scale field.

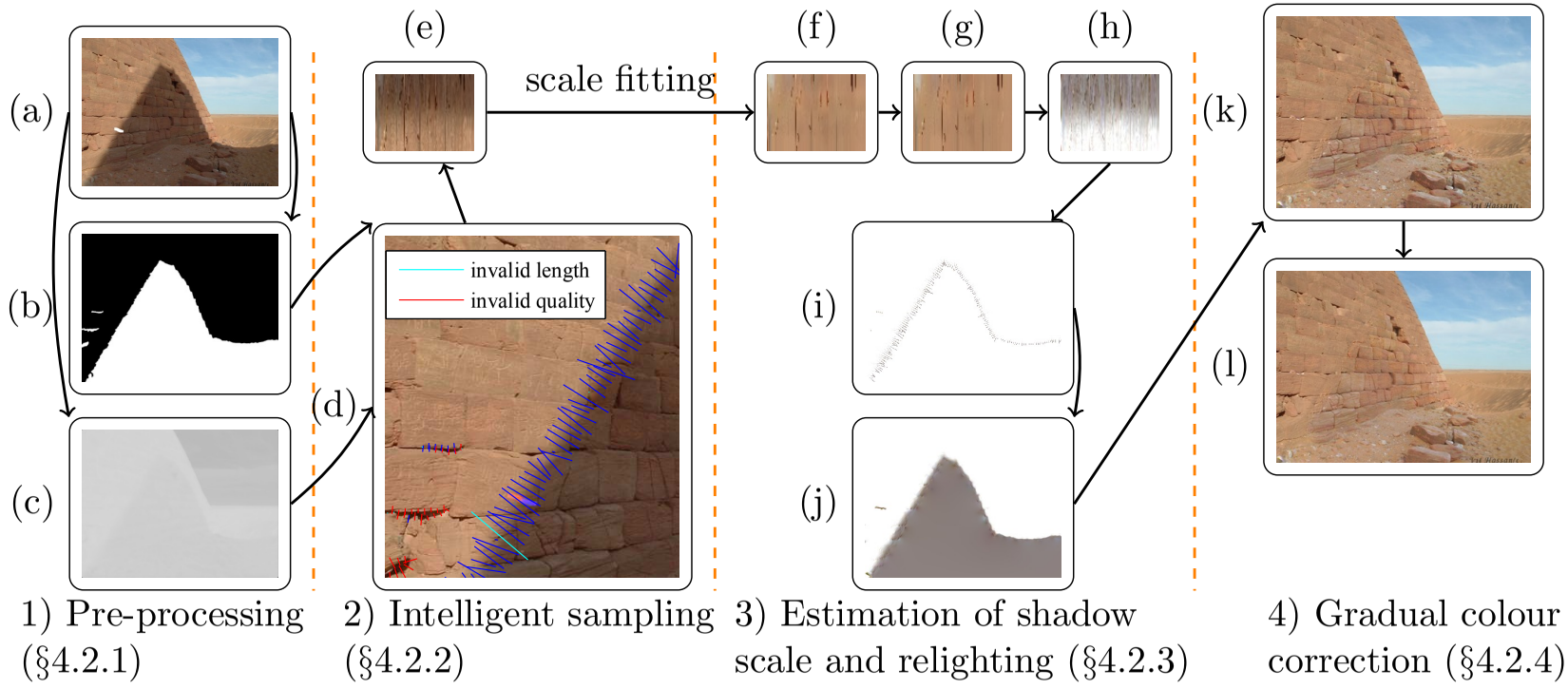


Figure 4-1: *Shadow removal pipeline. (a) input: a shadow image and user strokes in white covering both shadowed and lit pixels; (b) detected shadow mask; (c) fusion image; (d) initial penumbra sampling: the blue lines are valid samples, the other lines are invalid samples. The original single stroke has been divided into lit (red) and shadow (blue) parts; (e) boundary image of samples; (f) re-lit and filtered boundary image; (g) re-lit boundary image after boundary artefact removal; (h) rectified shadow scale of boundary image (each column refers to the scales for a sample); (i) sparse shadow scales; (j) dense shadow scales; (k) initial shadow-free image; (l) colour-corrected shadow-free image.*

4.2.1 Pre-Processing

Pre-processing provides a shadow mask and a fusion image (similar to the fusion image in Chapter 3) to assist intensity sampling (§4.2.2). Determining the initial shadow mask is the first step of shadow removal and is required in many previous methods including (Arbel and Hel-Or, 2011; Guo et al., 2012; Liu and Gleicher, 2008). Although some methods can achieve automatic shadow detection, these results are dependent on the quality and variation of training data. In this work, all that is required is the user to supply single strokes covering related shadow and lit pixels (Fig. 4-1(a)) – the remaining differentiation and recognition is fully automatic. Our interaction is easy to perform as it does not require users to explicitly distinguish between shadow and lit pixels nor to mark every shadow region (the interaction in Chapter 3). The pixels covered by the single user stroke are first classified as either shadow or lit pixels using a K-Means clustering (Seber, 2009). The feature used for clustering is the normalised RGB intensity and the normalised pixels coordinates. The cluster with the lowest mean for its RGB intensity is considered as a shadow cluster and vice versa. The classified input pixels' RGB intensities are used as the training features to construct a KNN classifier (Friedman et al., 1977) (number of neighbours: 3). Euclidean distance is used as the distance measure and the majority rule with nearest point tie-break as the classification measure. The input image can be binarised as a shadow mask, e.g. Fig. 4-1(b), using the pixel-wise KNN classifier.

The fusion image provides an illumination-insensitive layer, e.g. Fig. 4-1(c). It can be obtained by linearly fusing the channels of YCrCb colour space. The fused image $\mathcal{F}_p$ is computed as a weighted sum of the 3 normalised channels $\mathcal{C}_l$ as follows:

$$\mathcal{F}_p = (\sum_{l=1}^{3} \mathcal{C}_l \varphi(\sigma_l))/(\sum_{l=1}^{3} \varphi(\sigma_l)) \tag{4.1}$$

where l is the channel index, σ_l is the standard derivation of the umbra sample intensities of $\mathcal{C}_l$. φ is an exponential incentive function for determining the weight for each channel.

$$\varphi(x) = x^{-\lambda} \text{ subject to } \lambda > 0 \tag{4.2}$$

where x is the pixel intensity, λ (default value 5) determines the steepness of the incentive function. Lower variation of intensity is preferred as it means higher intensity uniformity in the umbra segment. To suppress texture noise, a median filter (Cliffs, 1990) with a 10-by-10 neighbourhood is further applied to $\mathcal{F}_p$.

4.2.2 Intensity Sampling

Our intensity sampling rejects inferior intensity samples for robust shadow scale estimation. There are three steps:

1) Adaptive raw intensity sampling RGB intensity profiles are extracted along sampling lines perpendicular to the shadow boundary, e.g. Fig. 4-1(d), where the boundary is obtained from the shadow mask. To accelerate shadow scale estimation, sampling lines are not measured at each shadow boundary point. Sparse and fixed distance sampling intervals are also avoided, as this may cause artefacts at highly uneven boundary segments (Arbel and Hel-Or, 2011; Liu and Gleicher, 2008). Instead, we adopt the interval-variable sampling used in our texture preserving shadow removal (§ 3.2.2).

To adapt the variance of penumbra softness, the length of a sampling line is guided by the fusion image. This problem is equivalent to finding the locations of the two ends of a sampling line. A bi-directional search is applied from each boundary point towards the lit area (end point) and the shadow area (start point) as described in Algorithm 1. The start and end points are initially set as the boundary point $\mathbf{p}_b$. To get the position

Algorithm 1: Sample end-point selection.

input : boundary point $\mathbf{p}_b$, boundary normal $\mathbf{n}_b$, fusion image $\mathcal{F}_p$
output: two ends ($\mathbf{p}_s$, $\mathbf{p}_e$) of a sampling line
$\widetilde{F} \longleftarrow \nabla\mathcal{F}_p$; $\mathcal{L} \longleftarrow \widetilde{F}(\mathbf{p}_b) \cdot \mathbf{n}_b$;
$\mathbf{p}_s \longleftarrow (\mathbf{p}_b)$; $\mathbf{p}_e \longleftarrow (\mathbf{p}_b)$;
repeat
  $\mathbf{v}_s \longleftarrow \widetilde{F}([\mathbf{p}_s])$; $\mathbf{v}_e \longleftarrow \widetilde{F}([\mathbf{p}_e])$;
  $\mathcal{L}_s \longleftarrow \mathbf{v}_s \cdot \mathbf{n}_b$; $\mathcal{L}_e \longleftarrow \mathbf{v}_e \cdot \mathbf{n}_b$;
  $\mathbf{p}_s \longleftarrow \mathbf{p}_s - \mathbf{n}_b$; $\mathbf{p}_e \longleftarrow \mathbf{p}_e + \mathbf{n}_b$;
until *either* $\mathbf{p}_s$ *or* $\mathbf{p}_e$ *is outside* $\mathcal{F}_p$ ***or*** $5(\mathcal{L}_s + \mathcal{L}_e) < \mathcal{L}$;

for a start point, the boundary normal $\mathbf{n}_b$ is iteratively subtracted from the start point (vice versa for the end point) until the average of two ends' projected gradient strength ($\mathcal{L}_s$ and $\mathcal{L}_e$) is small enough or either of the ends is the outside the range of image coordinates.

2) Sample selection Outlier intensity samples, e.g. Fig. 4-1(d), can affect accurate shadow scale estimation and cause unnatural shadow removal results. Unlike the slow and rough sample filtering in Chapter 3, two new criteria are adopted for this fast and reliable outlier detection. The first criterion is length of sampling line. The minimum length of a sampling line is 3 and the maximum length is $l_\mu + 3l_\sigma$ where l_μ and l_σ are the mean and the standard derivation of sample length respectively. The samples whose lengths are out of this range are removed. The second criterion is similarity of illumination change. A rough RGB intensity profile is obtained by down-sampling each intensity sample to 3 pixels' long using a discrete cosine transform. This intensity profile is then converted to the Log-domain. The approximate derivatives for each channel of each Log-domain intensity sample are supplied as illumination change features. A density-based DBSCAN clustering method (Ester et al., 1996) (number of neighbours:

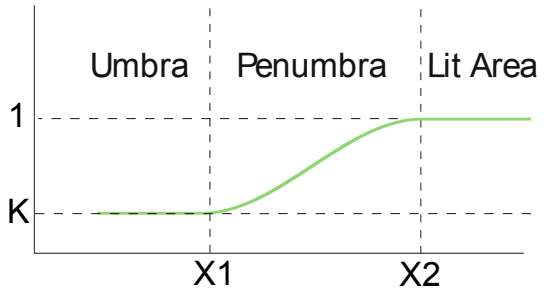


Figure 4-2: *Shadow scale model.*

3, radius: 0.2) is used to categorise samples. The samples belong to the largest cluster are identified as the samples with valid illumination change and the rest invalid samples are removed.

3) Sample de-noising Texture noise can still affect the smoothness of shadow scale estimation even when the outliers are removed. Our solutions in Chapter 3 does not handle this problem and some artefacts can be produced in the penumbra. In this algorithm, texture noise is removed from the selected intensity samples using a directional (i.e. parallel to normals of the shadow boundary) bilateral filter (Paris and Durand, 2009). To achieve this, the raw intensity samples are first resized as individual columns and their lengths are set as the maximum length of all raw samples. These columns are concatenated horizontally to form a boundary image, e.g. Fig. 4-1(e). A bilateral filter (Paris and Durand, 2009) is applied to each RGB channel of this image to suppress texture noise.

4.2.3 Estimation of Shadow Scale and Relighting

This sub-section explains the procedure for removing shadows based on the processed intensity samples. The following description of the algorithm is applied to the samples of each detected shadow boundary. There are three steps:

1) Initial intensity fitting Having obtained filtered and resized intensity samples at different positions along the boundary, our goal is to find illumination scaling values inside the umbra, penumbra and lit area. Different to the illumination change model adopted in § 3.2.2, the shadow scale change function S for each RGB channel of each intensity sample is modelled as follows (see also Fig. 4-2):

$$S(x) = \begin{cases} K & 0 \leq x \leq x_1 \\ f(x) & x_1 < x \leq x_2 \\ 1 & x_2 < x \leq 1 \end{cases} \tag{4.3}$$

where x is a normalised pixel location within the sampling line, x_1 and x_2 determine the start and end of the penumbra area respectively, and K is a positive scale constant for sample points within the umbra area ($x < x_1$). The constant 1 is assumed for the

lit area piece ($x > x_2$) as this falls inside a lit area of the image and does not require re-scaling. The function f is parametrised by K, v_1 and v_2 as follows:

$$\begin{cases} f(x) = (1-K)B(v_1(x-v_2)) + K \\ B(y) = -2y^3 + 1.5y - 0.5 \\ \begin{bmatrix} x_1 & x_2 \end{bmatrix} = v_2 + v_1^{-1} \begin{bmatrix} -0.5 & 0.5 \end{bmatrix} \end{cases} \tag{4.4}$$

where B is a cubic shape function (a sinusoidal function here also produces adequate results) and y is the input. Illumination of each channel may vary differently while penumbra boundaries of 3 channels are usually the same. A common penumbra width and position is thus assumed for each channel, determined by v_1 and v_2 respectively. Therefore, for each sampling line, there are 5 parameters to solve in total. The parameters can be solved by least squares fitting with a sequential quadratic programming algorithm (Nocedal and Wright, 2006). The fitted parameters of all sampling lines are also processed using a robust smoothing method (Garcia, 2010).

2) Boundary artefact removal For some over-processed images, directly relighting the sparse samples using the fitted shadow scales may cause band-like artefacts in the penumbra. Directional scale suppression (Fig. 4-1(f-h)) is therefore applied. Fig. 4-3 shows an example of more obvious boundary artefact removal.

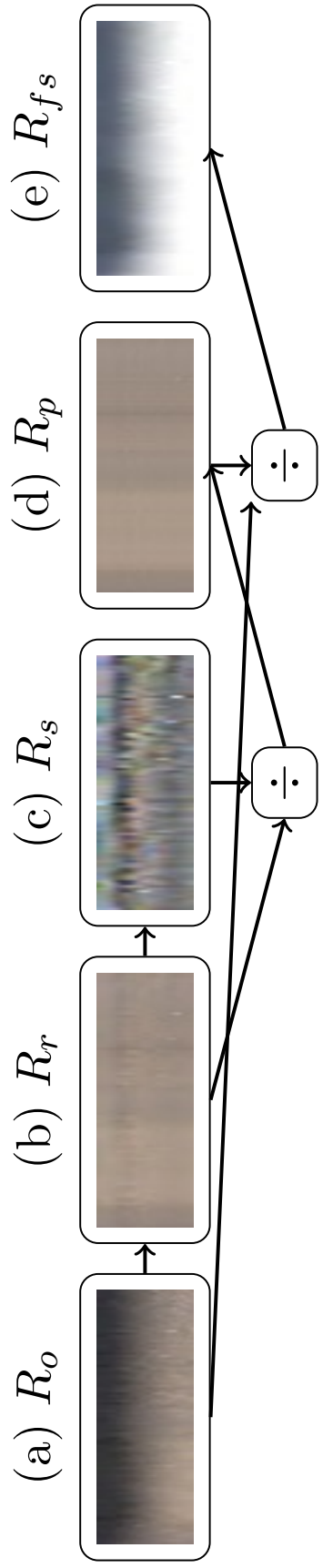


Figure 4-3: *Boundary artefact removal pipeline. The boundary image R_o is initially re-lit as R_r.. The band-like artefacts are R_r is extracted as R_s which is visualised by converting R_s to the log domain. R_p and R_{fs} are the processed artefact-free image and its shadow scale field respectively. The shadow removal results is shown in Fig. 4-5(b)*

.

The band artefacts appear as a pattern in each re-lit intensity profile (Fig. 4-3(b)). To suppress these, the previously re-sized and filtered intensity profile is first aligned according to the estimated penumbra position (x_1 and x_2) so that the illumination change of each resized sample is synchronised. The concatenated boundary image is denoted as R_o. The boundary image R_o is re-lit by inverse scaling according to Eq. 2.4 and the resulting image is termed R_r. It is assumed that band-like artefacts are locally similar and a local group size of 5 is specified for suppression. To extract the local band pattern, horizontal filtering is applied to R_r using an average kernel (size: 5x1) and the filtered image denoted as R_f. To suppress the band-like artefact, the variance of each column's, i.e. each samples, intensities of R_f are minimised. To achieve this, a variance image R_s is computed by dividing each column's intensities of R_f by its corresponding average intensity of that column. The artefact-free and re-lit boundary image is computed as $R_p = R_r \oslash R_s$ where $\oslash$ is an operator for element-wise division. Finally, the rectified scale image is computed as $R_{fs} = R_o \oslash R_p$. The sparse scales for each sampling lines corresponds to each column of R_{fs}.

3) Relighting To obtain a dense scale field (e.g. Fig. 4-1(j)), the sparse scales in the penumbra region are processed by smoothly interpolating and extrapolating the scales in other regions using spring-metaphor based in-painting (Bertalmio et al., 2000). The shadow-free image (e.g. Fig. 4-12(k)) can be obtained by inverse scaling according to Eq. 2.4.

4.2.4 Gradual Colour Correction

Some input images may have been significantly post-processed, e.g. through JPEG compression or gamma correction. Highly visible artefacts, e.g. differences in tone and contrast, may appear in shadow corrected areas as Eq. 2.4 does not hold in such cases. To address this, a simple gradual colour correction is introduced which is generally compatible for unknown post-processing affects. This step is only necessary for over post-processed images and the difference may otherwise be insignificant for the other images. The shadow removed image is first converted to L*a*b* colour space. It is assumed that the L*a*b* intensity variation of lower frequency is accurate and the errors appear in the intensity variation of higher frequencies. Fig. 4-4 shows the intermediate steps of colour correction in a* channel corresponding to the result in Fig. 4-1(l). Statistics are collected from the lit side pixels $\mathbf{P}_l$ and the umbra side pixels $\mathbf{P}_u$ both near penumbra as the target and source of colour correction respectively. In L*a*b* colour space, the image of higher frequency intensity variation $I^h = I^r - I^l$ is computed where I^l is the initial shadow removed image I^r filtered by a bilateral filter (Paris and Durand, 2009). The adjustment is completed in L*a*b* colour space as described in

Eq. 4.5.

$$\begin{cases} r_\sigma = \varsigma(I_c^h(\mathbf{P}_l))/\varsigma(I_c^h(\mathbf{P}_u)) \\ I_c^{ra}(\mathbf{P}_s) = I_c^l(\mathbf{P}_s) + r_\sigma I^h(P_s) \end{cases} \tag{4.5}$$

where c is the channel index, I^{ra} is the colour corrected image and the intensities of other unmodified pixel of I^{ra} are identical to those of I^r, ς is a function computes the median absolute deviation. Finally, to smooth the colour correction result, alpha blending is applied in RGB colour space according to the shadow scale as shown in Eq. 4.6.

$$I_c^f = I_c^r \circ \mathcal{S}_c^w + I_c^{ra} \circ (\mathbb{1} - \mathcal{S}_c^w) \tag{4.6}$$

where c is the channel index, $\mathcal{S}^w$ is the normalised scale field of $\mathcal{S}$, I_c^f is the final shadow-free image.

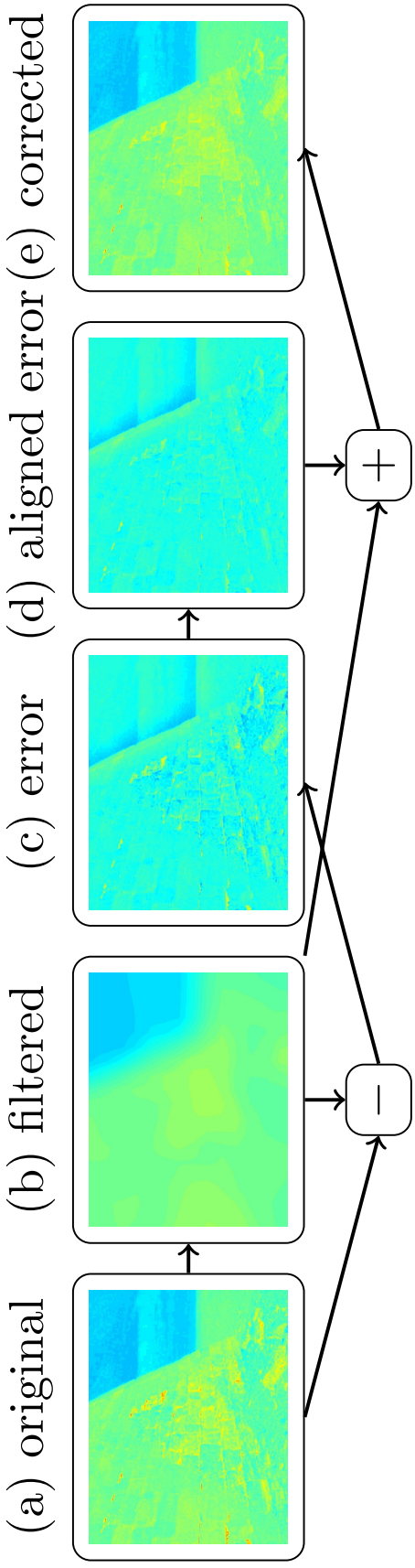


Figure 4-4: *Gradual colour correction pipeline. The a^* channel of initial shadow removed image (a) exhibits inconsistency between lit and shadow areas. The higher-frequency variation (c) of shadow and lit areas are aligned in (d). The corrected channel can be achieved by adding (b) and (d).*

4.3 Evaluation

In this section, technical improvements over our previous texture-preserving shadow removal (Chapter 3) are first described. We then show results of tests highlighting algorithm behaviour given variable user inputs. Finally, our algorithm is evaluated versus other state-of-the-art shadow removal methods using both visual comparisons and our improved quantitative evaluation measurements. This includes an additional algorithm for ground truth rectification on a current popular open data set (Guo et al., 2012).

4.3.1 Technical Improvements over Chapter 3

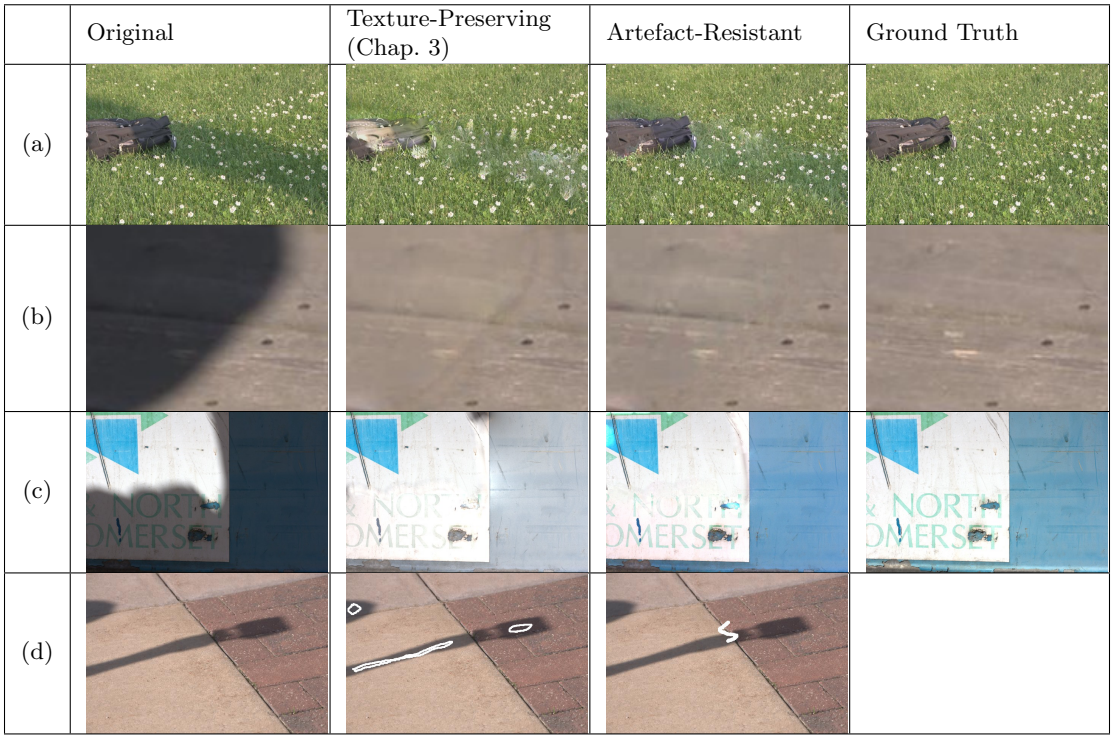


Figure 4-5: *Examples of improvements over Chapter 3. (a) intelligent sample selection (§4.2.2); (b) boundary artefact removal (§4.2.3); (c) chromatically-safe gradual colour correction (§4.2.4); (d) user input (in white).*

Compared with our previous texture-preserving shadow removal (Chapter 3), the following new improvements have been made:

1) An intelligent sample selection and de-noising As the sample selection in Chapter 3 is based on fitting error, outliers of incorrect illumination change may still be selected. The new sample selection method presented here avoids highly visible artefacts due to this (e.g. Fig. 4-5(a)).

2) More efficient scale estimation and a new boundary artefact remover Some band-like boundary artefacts may appear as our illumination change model may not

perfectly hold. The new boundary artefact remover presented here recovers a smoother penumbra and preserves its texture (e.g. Fig. 4-5(b)).

3) Improved colour correction The colour correction in Chapter 3 directly aligns the colour statistics near the penumbra which causes incorrect alignment when the colour distributions of lit and shadow sides are dissimilar. The improved colour correction method presented here resolves this issue (e.g. Fig. 4-5(c)).

4) Simplified user interaction Our new user interaction and shadow detection allow users to supply less accurate and fewer input strokes for complex shadow scenes (e.g. Fig. 4-5(d)). This new user interaction is compatible with the shadow removal solution in Chapter 3 as it also requires a binary shadow mask.

5) Quantitative evaluation The new work presented here provides a quantitative evaluation against other state-of-art methods, and includes comparisons using standardised bench-marks. Ground truth rectification on this bench-mark has also been employed to ensure a more accurate reflection of algorithm performance across methods.

4.3.2 Variability under Different User Inputs

Given user-specified single strokes, our shadow detection generates stable results in different conditions (e.g. Fig. 4-6(a) and Fig. 4-6(b)). In some cases, e.g. where the surface colour is very shadow-like, the detection results can be improved by supplying more than one stroke (e.g. Fig. 4-6(c) and Fig. 4-6(d)).

(a) single stroke test 1 (b) single stroke test 2

(c) result with 1 stroke (d) result with more strokes

Figure 4-6: *Variable input behaviours: The top row shows two examples using single input strokes. We supply 10 examples of single strokes placed in different locations and used as input (shown in red). The 2 grey-level images show the visualised probability of each pixel being marked in these 10 independent tests. Fewer grey pixels indicate higher stability, i.e. the image should only show black (0% probability) and white (100% probability) pixels when it is absolute stable. The bottom row shows examples highlighting how additional strokes can improve the detection result (binary mask).*

4.3.3 Rectification of Ground Truth

In the data set of (Guo et al., 2012), many of the shadow-free ground truth images are collected by entirely blocking the natural light in the scene. This unfortunately causes inconsistency in the brightness between some shadow-test images (e.g. Fig. 4-7(a)) and corresponding shadow-free ground truth (e.g. Fig. 4-7(b)). This will result in unfaithful quantitative evaluations in some test cases. To compensate for this, ground truth images of this kind can be globally re-lit (e.g. Fig. 4-7(c)) before evaluation. The RGB scale vector for global relighting can be estimated from the average RGB intensity of the common lit area. Lit pixels are first detected using a ratio image $I_{gr} = I \oslash I_{gt}$ where I is the original shadow image, I_{gt} is the shadow-free ground truth, $\oslash$ is an operator for element-wise division. K-Means clustering (Seber, 2009) is then used to divide the ratios into two clusters and the cluster with higher average ratios are identified as the lit cluster.

(a) shadow image (b) original GT (Guo et al., 2012) (c) our rectified GT

Figure 4-7: *Ground truth adjustment: An example rectified shadow-free ground truth (c) image obtain by correcting (b) shows a higher consistency with the test image (a). Note that in the original ground truth, the corrected image shows dark pixels as opposed to expected light ones (corrected in our rectified example).*

4.3.4 Quantitative Evaluation

In previous work (Guo et al., 2012), the quality of shadow removal is measured by the per-pixel Root-Mean-Square-Error (RMSE) between the shadow removal result and shadow-free ground truth in RGB colour space. However, the size and darkness of a shadow are often variable and this can result in biased shadow removal quality measurements. For example, an unprocessed image with a small area of shadow can have a smaller RMSE than the error of an image which has a large area of shadow that has only been partially corrected. Therefore, images with larger or darker shadows can affect the overall score. Our error ratio $\mathbf{E}_r$ is therefore computed as follows:

$$\mathbf{E}_r = \mathbf{E}_n / \mathbf{E}_o \tag{4.7}$$

where $\mathbf{E}_n$ is the RMSE between the shadow-free ground truth and shadow removal result, and $\mathbf{E}_o$ is the RMSE between the shadow-free ground truth and the original shadow image. All error measurements are assessed in RGB colour space. This normalised measure reflects the degree of shadow removal towards the ground truth independent of original shadow intensity and size. To assess the robustness, the standard derivation is also computed for each measurement.

Extending previous work on ground truth based evaluation (Guo et al., 2012), the *category* of shadow is also considered in our work for the first time, i.e. the type of shadow being cast. In our test, additional scores for particular categories of shadows with soft penumbra and strong texture background are shown. These special categories are included because they are generally more difficult to process compared with low texture backgrounds with compact shadows. Quantitative results are presented in Tab. 4.1, where *starred* columns refer to pixels just in the shadow region being considered, and *non-starred* columns refer to the entire image. The test results are divided into 2 sub-tables: the top sub-table shows the result for all test cases and the bottom sub-table shows the result only for test cases without significant detection failures which therefore gives an indication of shadow *removal* performance independent of shadow *detection* quality. Our method outperforms the other approaches compared against for all of the scores. There are a small number of our scores that are numerically close to the second best ones. However, small numerical differences may indicate visually significant artefacts which are shown in the visual comparison sub-section (§4.3.5).

	Yang et al. (2012)		Guo et al. (2012)		Su and Chen (2010)		Ours	
	RMSE	RMSE*	RMSE	RMSE*	RMSE	RMSE*	RMSE	RMSE*
Soft	1.45 (1.15)	1.23 (1.32)	0.79 (0.40)	0.62 (0.37)	0.65 (0.31)	0.69 (0.66)	**0.51 (0.29)**	**0.34 (0.29)**
Texture	1.14 (0.30)	0.86 (0.29)	0.72 (0.41)	0.46 (0.36)	0.77 (0.35)	0.61 (0.42)	**0.62 (0.28)**	**0.39 (0.23)**
All	1.31 (0.68)	1.19 (0.95)	0.67 (0.35)	0.48 (0.36)	0.79 (0.42)	0.97 (1.13)	**0.59 (0.28)**	**0.40 (0.31)**

	Yang et al. (2012)		Guo et al. (2012)		Su and Chen (2010)		Ours	
	RMSE	RMSE*	RMSE	RMSE*	RMSE	RMSE*	RMSE	RMSE*
Soft	1.52 (1.37)	1.35 (1.59)	0.68 (0.47)	0.44 (0.34)	0.66 (0.35)	0.67 (0.70)	**0.49 (0.27)**	**0.30 (0.27)**
Texture	1.15 (0.31)	0.86 (0.30)	0.69 (0.42)	0.40 (0.34)	0.81 (0.35)	0.64 (0.43)	**0.65 (0.28)**	**0.40 (0.23)**
All	1.31 (0.71)	1.19 (0.99)	0.62 (0.35)	0.39 (0.33)	0.81 (0.45)	1.01 (1.21)	**0.59 (0.28)**	**0.39 (0.30)**

Table 4.1: *Categorised quantitative test results (top: results for all (79) test cases; bottom: results for 67 test cases without significant detection failures for all methods tested). The non-stared and stared columns indicate the error score where all pixels in the image are used, and just shadow area pixels respectively. The categories of "Soft" and "Texture" refer to shadow images with soft penumbra and strong texture background respectively. "All' is the average result for all test cases in the data set, including those categorised as "Soft" and "Texture". Method (Guo et al., 2012) is trained using a large shadow detection data set from (Zhu et al., 2010). The best score for each measurement is in bold and the second best score is underlined. The standard deviation for each measurement are in brackets. In the our ordering, the average error is compared before comparing the standard derivation.*

4.3.5 Visual Comparisons

Figure 4-8: *Comparisons of our method against existing state-of-the-art techniques using ground truth images from (Guo et al., 2012). Please magnify to examine in detail.*

Typical examples of our shadow removal algorithm are shown for visual comparison in Fig. 4-8. Overall, our method produces more qualitatively pleasing removal results against the evaluated methods. However, minor artefacts are sometimes noticeable when the input image has a highly irregular soft penumbra, or the background of the shadow area is highly textured. The method of (Su and Chen, 2010) shows better overall visual results compared with (Guo et al., 2012) but the approach over-corrects global illumination causing lower numerical scores in Fig. 4-9. Fig. 4-9(a-b) show two failure cases where shadows are soft and overlapped, and there are some small visible

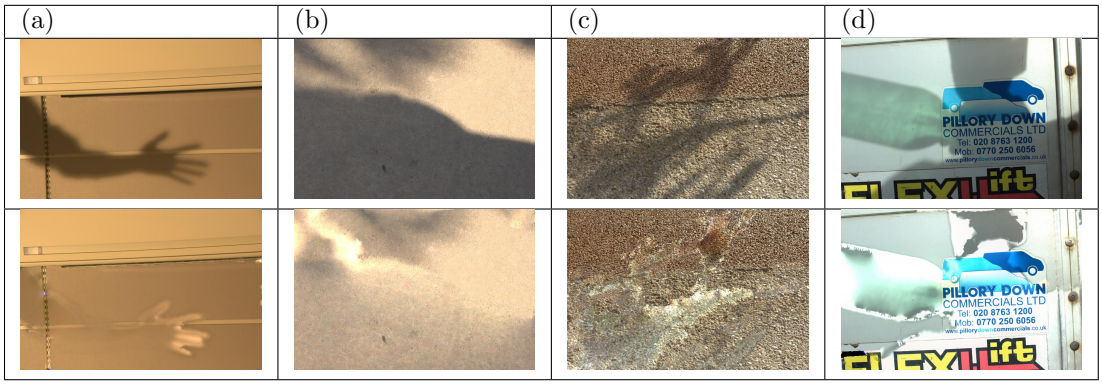


Figure 4-9: *Failure cases – all the shadow removal methods tested in this chapter create highly visible. These cases present future research challenges.*

artefacts across all methods. Fig. 4-9(c) highlights a failure case where the shadow is highly broken. Fig. 4-9(d) shows a failure case where the shadow area illumination is multi-coloured, i.e., a combination of greenish and bluish shadows. These examples of limitation cases identified in all the tested shadow removal methods represent the current research challenges in our field.

4.4 Summary

A user-friendly shadow removal method has been presented which provides several improvements over the shadow removal solution in Chapter 3. This includes – easy user input, intelligent intensity sampling, a local group processing based shadow scale estimation and robust colour correction. Unlike the visual inspection only evaluation in Chapter 3, the presented algorithm in this chapter has been quantitatively evaluated using a standard data set in this area, and demonstrates state-of-the-art performance. Visual comparisons are also presented for a large number of shadow removal cases taken from the evaluation data set. Through our analysis, difficult shadow removal cases such as highly-soft shadows and shadows on strong texture background have also been identified. These represent exciting research challenges in our area, and highlight the fact that single image shadow removal still represents an open problem in computer vision. In Chapter 6, we present a fast shadow removal method with some significant improvements in processing speed and the shadow removal quality for more difficult cases which include coloured shadows and highly broken shadows. This is verified by a new large, reliable, and diverse ground truth data set for shadow removal presented in the following Chapter 5.

Publications

The following publication is related to this chapter:

Gong, H. and Cosker, D. (2014d). User-assisted image shadow removal. *Computer*

Vision and Image Understanding (CVIU). under review

Chapter 5

Shadow Removal Ground Truth for Variable Scene Categories

Overview

In this chapter, we propose a new data set for shadow removal evaluation which overcomes several issues in the previous state-of-the-art data set in terms of robustness and variability. We first give an introduction of our data set as well as an overview of the motivations and difficulties for creating such a data set. We then describe the content of the data set and the procedure for reliable capturing and assessment for quality of ground truth. Finally, we describe our on-line benchmark for shadow removal evaluation.

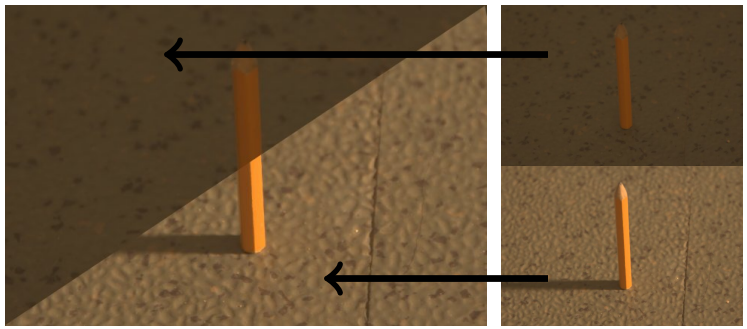
5.1 Introduction

A thorough quantitative evaluation of shadow removal performance requires a high-quality, diverse shadow-free ground truth. The first public data set was supplied by Guo et al.. However, while this is a valuable resource for evaluating shadow removal and the first of its kind, there are many opportunities for expansion and several improvements are presented in our new data set. As mentioned previously in the background chapter (Chapter 2), it is desirable to have the shadow removal methods quantitatively evaluated based on a reliable and large data set. In Chapter 4, we attempt to improve the only existing data set (Guo et al., 2012) by artificially amending the problematic ground truth data. However, it is still not satisfying due to the following issues:

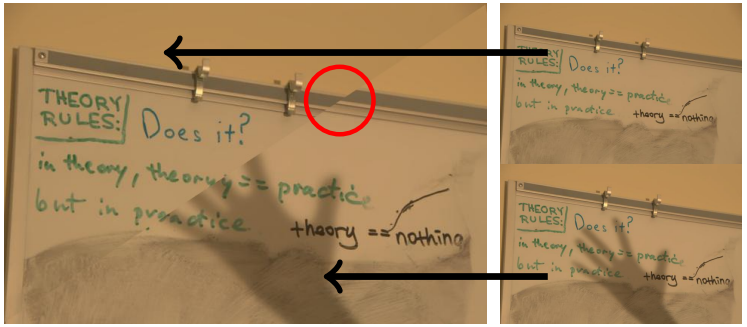
- The amended data is not naturally captured.
- Significant image registration errors still exist in many of their test cases.

- The data set only contains a relatively small number of test cases and the images are not labelled according to their appearance attributes of shadow.

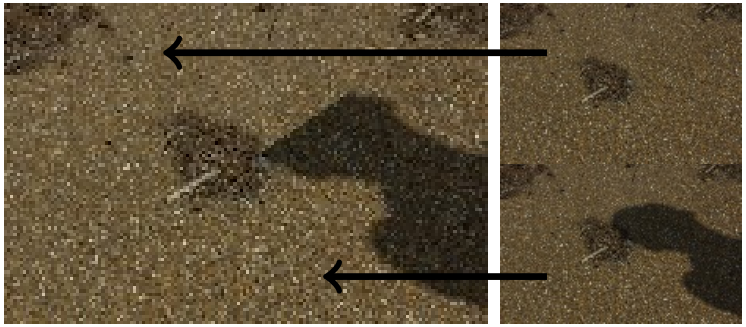
It is therefore necessary to build a new data set for comprehensive shadow removal evaluation. Firstly, the concept of shadow categories is introduced for the first time in our proposed data set, and a wide range of new types proposed. Secondly, ground truth is constructed and verified in a careful manner so as to remove irregularities between test and validation images. In terms of the latter, we note environmental illumination and registration errors between some shadow and ground truth images in existing data sets. An example of comparison is shown in Fig. 5-1. Our new data set avoids these issues and represents, we believe, the most stable and thorough data set for shadow removal evaluation available today. In order to highlight the benefits of our rigorous data protocol, in § 5.4 the quality of our ground truth data is quantitatively compared to another state-of-the-art set (Guo et al., 2012).



(a) mismatched illumination in the lit area



(b) unregistered pixels between shadow image and its shadow-free ground truth image



(c) our data (no artefacts)

Figure 5-1: *Visual comparison of ground truth data for shadow removal. For each image: top left segment – shadow-free image; bottom right segment – shadow image. (a) and (b) are taken from (Guo et al., 2012) which reflect the two annotated issues. An example from our data – which rejects image pairs with these properties is shown in (c).*

5.2 Data Capture

Figure 5-2: *Ground truth capture rig. The capture is controlled by a wireless trigger and tripod is used to stabilised the rig. This setup minimises both registration error and illumination inconsistency between shadow and shadow-free images but they can not be avoided.*

Shadow images and their ground truth are captured using a camera with a tripod and a remote trigger (e.g. Fig. 5-2). This rig minimises misalignment due to camera shake. To minimise illumination variance, images are captured within a very short interval of time using a manual capture mode with fixed ISO and exposure settings. These capture settings may vary for different test cases but the capture settings for a shadow image and its corresponding shadow-free image are consistent. We use a occluder to create shadow images and move it away to capture shadow-free ground truth. When collecting data, environmental effects are often unavoidable, e.g., wind can cause camera shake or the sun might move behind the clouds. Such failed acquisitions are rejected from our data set using a quantitative assessment outlined in §5.4. For evaluation, our shadow data is also categorised according to 4 different attributes: degree of texture, shadow softness, brokenness of shadow, and colour variation. In each category, the shadow images are further labelled according to 3 intensity degrees. In total, our final data set after rejection consists of 186 test cases. For comparison, the previous state of the art from (Guo et al., 2012) consists only of 28 test cases after applying our strict rejection measure.

5.3 Data Content

This data set is made for robustly evaluating the quality of shadow removal for shadows created by a single light source. In our combined data set, 206 shadow images are captured in outdoor scenes and 8 shadow images are captured indoor. Fig. 5-3 shows the frequency statistics of the categories of shadow in this data set. For most categories, the number of test cases for weak intensities are significantly more than the number of the other stronger intensity levels. This is due to a shadow image of one kind containing the other shadow attributes. To minimise this effect of mixture, most shadow images are captured with the appearance of only one predominant shadow attribute. This is the reason why the test cases for weak intensities are significantly more. However, as long as there are sufficient test cases for an experiment and the scores are averaged, the result should still remain robust. The labels of each shadow image used in our evaluation are provided on our evaluation website (Section 5.5) as a MATLAB matrix file and a text file.

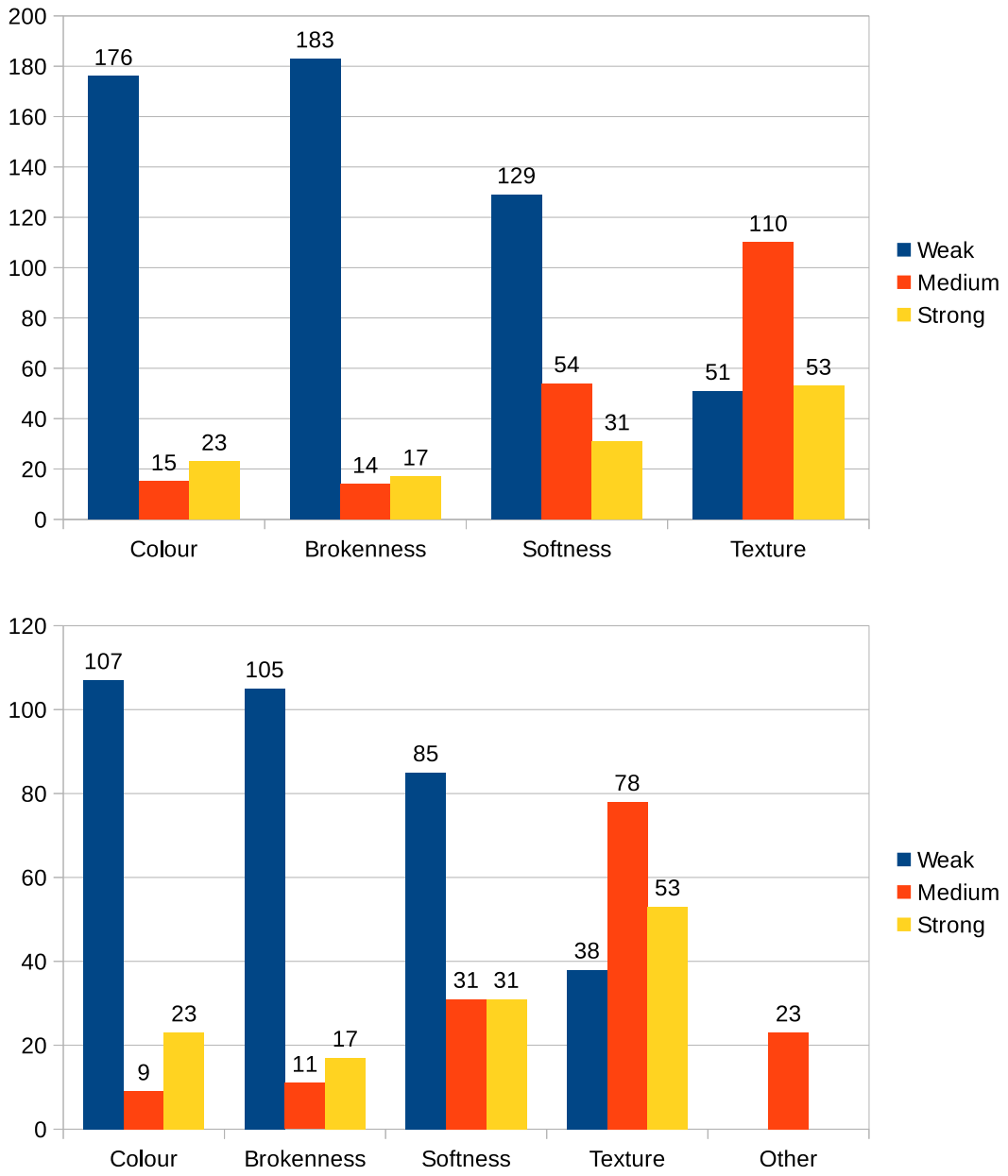


Figure 5-3: *Frequency of test cases in our shadow removal data set. Top: test case counts in each category and each intensity level. Bottom: test case count with extreme cases excluded. In the bottom chart, for each count of each attribute, the shadow images with other predominant attributes (degree = 3) are not used. Hence, test cases have a strong single bias towards one of the attributes. "Other" refers to a set of shadow cases showing no markedly predominant attributes (degree = 1).*

5.4 Quantitative Evaluation

Ideal pairs of ground truth images should have a minimum intensity difference in the common lit area – which will also indicate whether registration is poor (due to camera shake or scene movement – which should be rejected). This is utilised to assess the quality of ground truth candidates. The error image ΔI and the ratio image I_r are first computed as follows:

$$\Delta I = I_s - I_g \tag{5.1}$$

$$I_r = \Phi(I_s) \oslash \Phi(I_g) \tag{5.2}$$

where I_s and I_g are shadow and shadow-free images, $\oslash$ is element-wise division and Φ is a function that converts RGB image to grey-scale image. The set of pixels P_r of I_r that satisfies $I_r(P_r) \geq 1$ are regarded as lit pixels. The ground truth error Q_d is computed as follows:

$$Q_d = \mu(|\Delta I(P_r)|) + \sigma(\Delta I(P_r)) \tag{5.3}$$

Ground truth pairs in our data set with $Q_d > 0.05$ are removed. Using this measure, our initial data capture of 195 test cases results in 186 test cases with stable illumination changes between the shadow and ground truth images. Comparing to the quality of other ground truth data sets, (Guo et al., 2012) (state-of-the-art) results in mean error of 0.18 (leaving 28 out of 79 test cases) while ours is 0.02.

5.5 On-Line Shadow Removal Benchmark

We provide a website of on-line shadow removal benchmark for users to conveniently evaluate their new shadow removal algorithms. Similar to the alpha matting evaluation website (Rhemann et al., 2009), our shadow-free image data is also made hidden to prevent unfair tuning of the algorithm to the ground truth. To use this benchmark, users are required to process a large set (214 test cases) of shadow images using their own shadow removal algorithm. The processed shadow-free images are uploaded to our website. Our website then gives the detailed and categorised quantitative result for these processed images. Our website is accessible via `http://www.cs.bath.ac.uk/~hg299/shadow_eval/`. A screen-shot of our website is shown in Fig. 5-4.

5.6 Summary

A state-of-the-art data set for shadow removal evaluation and an on-line benchmark website for shadow removal based on this data set are proposed. We have compared our data set with the previous state-of-the-art data set and shown the significant im-

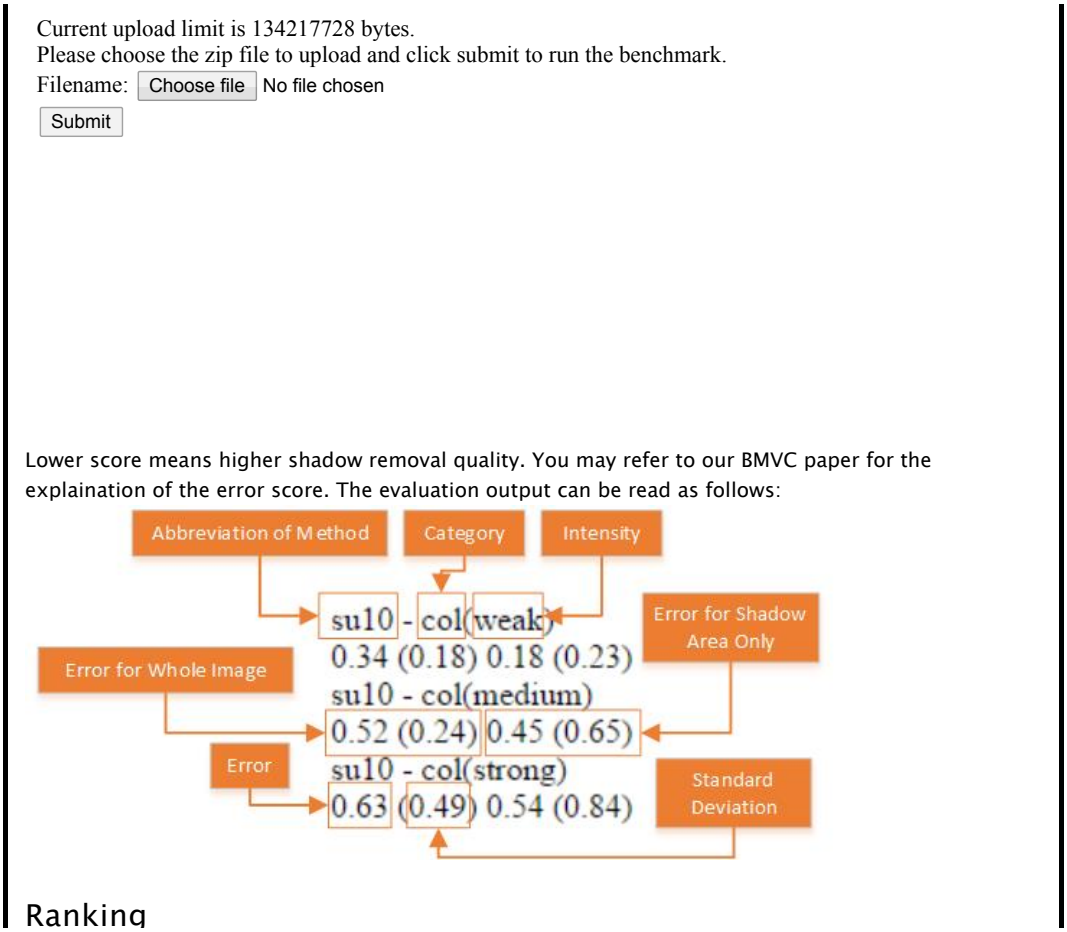


Figure 5-4: *A screen-shot of our website for on-line shadow removal benchmark.*

provements which are visually and quantitatively verified. In the following Chapter 6, this new data set is used to evaluate a new fast shadow removal method as well as the methods proposed in the previous chapters and some other state-of-the-art shadow removal methods.

Publications

The following publications are related to this chapter:

- Gong, H. and Cosker, D. (2014c). Interactive shadow removal and ground truth for variable scene categories. In *British Machine Vision Conference (BMVC)*. Best Student Paper
- Gong, H. and Cosker, D. (2014b). Interactive shadow removal and ground truth for variable scene categories. *International Journal of Computer Vision (IJCV)*. under revision

Chapter 6

Fast Shadow Removal

Overview

We present an interactive, robust and high quality method for fast shadow removal. To perform detection, we use an on-the-fly learning approach guided by two rough user inputs for the pixels of the shadow and the lit area. From this we derive a fusion image that magnifies shadow boundary intensity change due to illumination variation. After detection, we perform shadow removal by registering the penumbra to a normalised frame which allows us to efficiently estimate non-uniform shadow illumination changes, resulting in accurate and robust removal. For quantitative evaluation, we adopt our first reliable, validated and multi-scene category ground truth (Chapter 5) for shadow removal algorithms. Using our data, we perform the most thorough comparison of state-of-the-art shadow removal methods to date. The experimental result shows that our algorithm outperforms existing state-of-the-art methods.

6.1 Introduction

In this chapter, we propose an interactive, high-quality and robust method for fast shadow removal using two rough user-defined strokes indicating the shadow and lit image areas. Compared with the single stroke user interaction in Chapter 4, this user interaction may require users to supply a little more strokes. However it provides the additional support for the cases where it is difficult to cover two sets of related lit and shadow pixels in only one single stroke. Given detection, we deliver reliable shadow removal – verified with thorough quantitative tests for different types of shadow scene (for the first time in this area) comparing to previous state-of-the-art approaches. We also adopt our large high-quality and multi-scene category ground-truth data set, which is presented in Chapter 5, for the evaluation of shadow removal. This data set is quantitatively verified to ensure registration and illumination errors between test and ground truth images are eliminated. Our approach represents a state-of-the-art method

for shadow removal, with the most robust comparative evaluation of such approaches to date.

In this method, the shadow effect is considered as a Hadamard product of a shadow scale layer and a shadow-free image as introduced in Eq. 2.4.

6.1.1 Contributions

Given our overview of state-of-the-art approaches, we propose two main contributions:
1) Simple user input based shadow detection: Our shadow detection component requires only two rough user scribbles marking samples of lit and shadow pixels. Our approach differs from previous work and the previous two chapters (Chapter 3 and Chapter 4) requiring more complex user-inputs or simpler inputs that compromises robustness and quality and is more compatible with complex shadow scenes.
2) High quality and fast shadow removal: Unlike existing methods requiring slow pixel-wise optimisation or an inflexible fitting model (e.g. Chapter 3 and Chapter 4), we introduce penumbra unwrapping to derive a few unified intensity samples across the penumbra for robust and efficient estimation of illumination changes without requiring prior training. Our method is simple and fast yet offers state of art shadow removal quality.

To summarise, we believe our contributions are important to this area of research due to our significant improvements over the state-of-the-art in shadow removal in a wide range of thorough and repeatable tests.

6.2 Fast Shadow Removal

In this section, we overview our algorithm first in brief, and then expand on technical details for each of its components. Our algorithm consists of three steps (see Fig. 6-1):
1) Pre-processing (§6.2.1) We detect an initial shadow mask (Fig. 6-1(b)) using a KNN classifier trained from data from two rough user inputs (e.g. Fig. 6-1(a)). We generate a *fusion image*, which magnifies illumination discontinuities around shadow boundaries, by fusing channels of YCrCb colour space and suppressing texture (Fig. 6-1(c)).
2) Penumbra unwrapping (§6.2.2) Based on the detected shadow mask and fusion image, we sample the pixel intensities of sampling lines perpendicular to the shadow boundary (Fig. 6-1(d)), remove noisy ones and store the remaining as columns for the initial penumbra strip (Fig. 6-1(e)). We align the initial columns' illumination changes using its intensity conversion image (Fig. 6-1(f)). This results in an aligned penumbra strip (Fig. 6-1(g)) whose conversion image (Fig. 6-1(h)) exhibits a stabler profile.
3) Estimation of shadow scale and relighting (§6.2.3) From the penumbra strip, we synthesise a few unified samples, e.g. Fig. 6-1(i), of intensity change which allows

for a fast and robust estimation of illumination change along sampling lines to derive the sparse scales for all sampled sites (Fig. 6-1(j)) which are propagated to form a dense scale field (Fig. 6-1(k)). We remove shadows by inverse scaling using this non-uniform field (Fig. 6-1(l)).

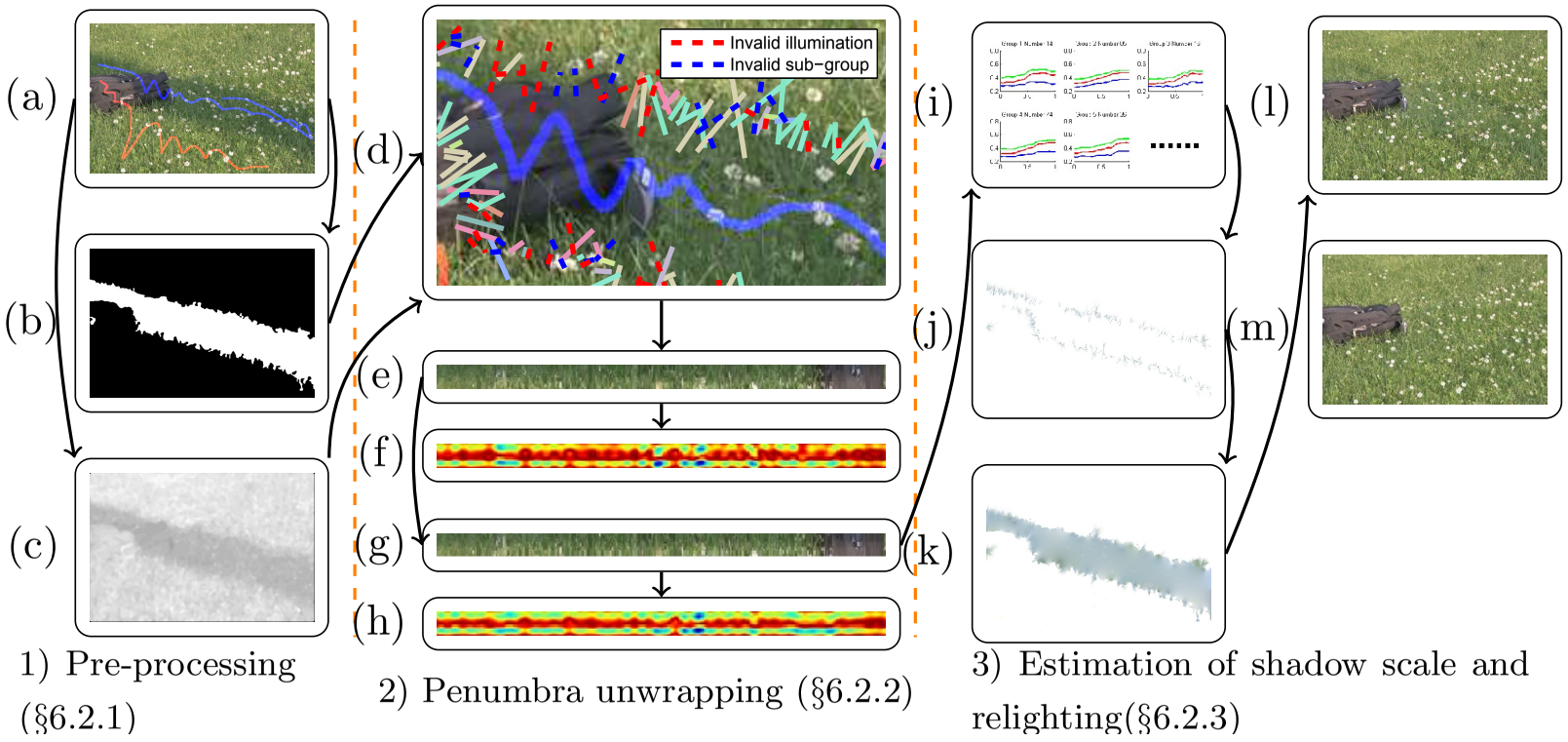


Figure 6-1: *Fast shadow removal pipeline. (a) input: a shadow image and user strokes (blue for lit pixels and red for shadowed pixels); (b) detected shadow mask; (c) fusion image; (d) initial penumbra sampling (solid lines in different colours indicate valid samples of different sub-groups. Dashed lines are invalid samples removed using two criteria); (e) initial penumbra regularisation; (f) initial penumbra conversion image; (g) final penumbra regularisation; (h) final penumbra conversion image; (i) penumbra illumination estimation; (j) sparse shadow scale; (k) dense shadow scale; (l) output; (m) ground truth.*

6.2.1 Pre-Processing

Pre-processing provides a detected shadow mask and a fusion image to assist penumbra unwrapping. Although there have been automatic methods for shadow detection, results are dependent on training data quality and variation. Instead, our method requires no prior training or learning – only two user-supplied rough inputs indicating sample lit and shadow pixels (Fig. 6-1(a)). We supply the marked pixels' RGB intensities in the Log domain as the training features and construct a KNN classifier. We choose the Euclidean distance as the distance measure and the majority rule with nearest point tie-break as the classification measure. We apply spatial filtering with a Gaussian kernel (size = 5, SD = 3) to the obtained image of posterior probability and binarised filtered image using a threshold of 0.5. Although detection errors along the boundary, as well as post-filtering, can result in intensity samples with unsynchronised illumination changes at sharp boundary, our penumbra unwrapping and alignment step (§6.2.2) can compensate for this. Thus, our shadow *removal* method is somewhat robust to noise in the initially detected shadow mask, and would also be applicable to alternative (e.g. automatic) detection methods.

To assist unwrapping of the penumbra, we derive an image that magnifies illumination discontinuities around the shadow boundary – also assisting penumbra location – which we call the *fusion image* (e.g. Fig. 6-1(c)). There are two steps:

1) Magnification of illumination discontinuity We derive an initial fusion image $\mathcal{F}$ that maximises the contrast between shadow and lit areas by linearly fusing the three channels ($\mathcal{C}_l$) of YCbCr space as follows:

$$\mathcal{F} = \sum_{l=1}^{3} a_l \mathcal{C}_l \text{ subject to } \sum_{l=1}^{3} a_l = 1 \tag{6.1}$$

where a_l is the fusing factor of $\mathcal{C}_l$ (positive). The best fusing factors are derived by minimising the following objective function E_b:

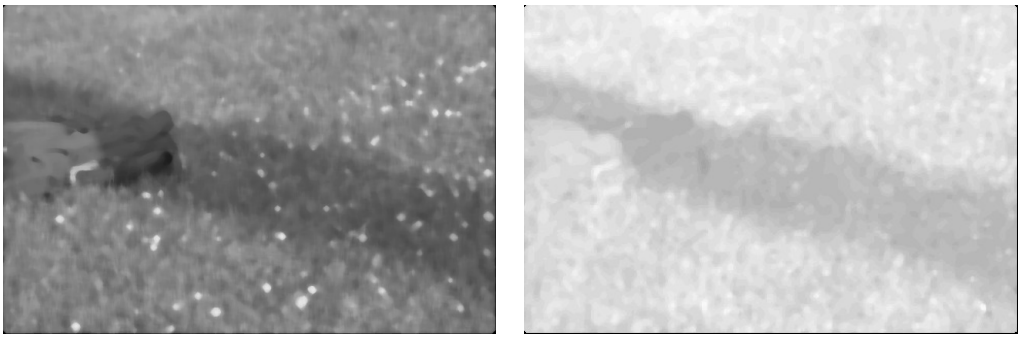
$$E_b(\mathbf{a}) = \frac{\mu(\mathcal{F}_\mathbf{S})}{\mu(\mathcal{F}_\mathbf{L})} + \frac{\sigma(\mathcal{F}_\mathbf{S}) + \sigma(\mathcal{F}_\mathbf{L})}{\sigma(\mathcal{F}_{\mathbf{S}\cup\mathbf{L}})} \tag{6.2}$$

where $\mathbf{a}$ is the vector of fusing factors and $\mathcal{F}_\mathbf{S}$ and $\mathcal{F}_\mathbf{L}$ are the two sets of shadow and lit pixels marked by user scribbles. In this chapter, we define σ and μ as functions that respectively compute the standard deviation and mean of a set of values. The first term ensures larger distinction between pixels of lit and shadow regions and the second term ensures smaller variation for pixels of the same lit or shadow regions.

2) Suppression of texture We reduce image texture by applying a median filter with a 10-by-10 neighbourhood to $\mathcal{F}$.

In our experiments, we found YCbCr colour space to offer more perceptually meaningful information, while illumination in RGB space can be affected by texture noise.

An example comparison of fusing channels in YCrCb colour space compared to RGB colour space is shown in Fig. 6-2.



(a) RGB fusion image (b) YCrCb fusion image

Figure 6-2: *Comparison of colour space for fusion: The same optimisation scheme is applied to the image in Fig. 6-1(a) but using different colour spaces. Texture noise, such as in the flowers, appears more pronounced in the RGB fusion image (left) whilst the YCrCb fusion image (right) presents cleaner illumination information.*

6.2.2 Penumbra Unwrapping

The shadow boundary generally has a noisy profile with variable penumbra width. This can lead to inaccurate estimation of shadow scales and resulting artefacts. We therefore unwrap the penumbra into a strip and align its sampled columns of illumination change. This improves the detection of outliers and allows linearisation of processing in the penumbra – leading to significant gains in efficiency and speed (see Fig. 6-1(g)).

We sample the intensity of sampling lines perpendicular to the shadow boundary (Fig. 6-1(d)) as columns for the initial penumbra strip. The length of a sampling line is determined by locating suitable start and end points guided by the fusion image $\mathcal{F}$. Similar to Algorithm 1, we start a bi-directional search from each boundary point that extends the sampling line towards the lit area (end point) and the shadow area (start point) as described in Algorithm 2. We initially set the start and end points as the boundary point (x_b, y_b) and the direction vector $\Delta\mathbf{v}$ as the normalised gradient vector of (x_b, y_b). Compared with Algorithm 1, the direction vector for gradient projection used here is the gradient vector at a boundary point rather than its normal vector. This is practically more reliable because the texture of the surface in the penumbra causes significant noise which is not useful for determining the termination of sampling. To get the position for a start point, we iteratively subtract $\Delta\mathbf{v}$ from the start point until its projected gradient is small enough (vice versa for the end point).

To avoid outliers, e.g. sampling lines at occlusion boundaries, we filter invalid samples based on an assumption of similar shadow scales. We first compute a scale vector $Y_c = T_l - T_s$ where T_l and T_s are the average Log-domain RGB intensities of

Algorithm 2: Sample end point selection

input : boundary point (x_b, y_b), fusion image $\mathcal{F}$
output: two ends ($\mathbf{p}_s$, $\mathbf{p}_e$) of a sampling line
$\widetilde{F} \longleftarrow \nabla\mathcal{F}$; $\mathcal{L} \longleftarrow |\widetilde{F}(x_b, y_b)|$; $\Delta\mathbf{v} \longleftarrow \widetilde{F}(x_b, y_b)/\mathcal{L}$;
$\mathbf{p}_s \longleftarrow (x_b, y_b)$; $\mathbf{p}_e \longleftarrow (x_b, y_b)$;
repeat
  $\mathbf{v}_s \longleftarrow \widetilde{F}([\mathbf{p}_s])$; $\mathbf{v}_e \longleftarrow \widetilde{F}([\mathbf{p}_e])$;
  $\mathcal{L}_s \longleftarrow \mathbf{v}_s \cdot \Delta\mathbf{v}$; $\mathcal{L}_e \longleftarrow \mathbf{v}_e \cdot \Delta\mathbf{v}$;
  $\mathbf{p}_s \longleftarrow \mathbf{p}_s - \Delta\mathbf{v}$; $\mathbf{p}_e \longleftarrow \mathbf{p}_e + \Delta\mathbf{v}$;
until $\mathbf{p}_s$ *or* $\mathbf{p}_e$ *is not within the range of* $\mathcal{F}$ ***or*** $10\mathcal{L}_s > \mathcal{L}$ *or* $10\mathcal{L}_e < \mathcal{L}$;

the lit and shadow halves of a sample. We convert Y_c to spherical coordinates as a feature vector Y_s. We apply DBSCAN clustering (Ester et al., 1996) (radius:0.2) to Y_s of all samples and store the samples that belong to the largest cluster as valid ones with valid illumination. For finer scale estimation, we divide the valid cluster into a few sub-groups using mean-shift (Comaniciu and Meer, 2002) (band width:0.06) and discard the samples of the invalid sub-groups whose numbers are less than 10% of the largest sub-group's. Fig. 6-1(i) shows an example of the above sample categorisation.

As the lengths of samples are different, we normalise this by re-sizing all the samples to their average length n_a. The normalised samples are concatenated as columns to form the initial penumbra strip. The vertical illumination change of columns at this stage may still be misaligned. We therefore resolve this using fine-scale alignment focusing on vertical centre adjustments and scaling of the columns. Fig. 6-1(g) shows an example of the aligned strip and Fig. 6-3 visualises the process of alignment. To assist alignment, we amplify illumination changes by deriving a fusion image $\mathcal{W}_p$ of the strip. Similar to our previous formulation, we first fuse the strip using Eq. 6.1 with factors $\mathbf{a}$ and apply an ideal low-pass frequency filter using a cut-off $k_w = \max([n_a/8], 2)/n_a$. To locate the centre, we generate a conversion image $\mathcal{O}(i) = -|\mathcal{W}_p(i) - \mu(\mathcal{W}_p(i))|$, e.g. Fig. 6-1(f), where i is the column number. In $\mathcal{O}$, the centres of the columns of strips appear as peaks. We can find the peaks of columns from $\mathcal{O}$ and vertically align based on these.

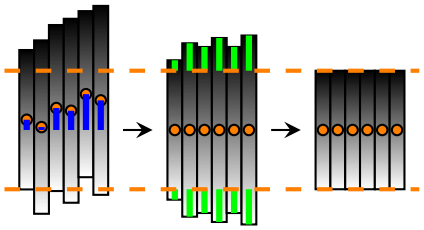


Figure 6-3: *Alignment of penumbra strip: The orange circles are the centres of columns in the penumbra strip. The orange dashed lines indicate the desired column length for a strip. The blue and green lines indicate the shifts required. The alignment is in two steps: (left to middle) alignment of centre; (middle to right) alignment of illumination change.*

To further ensure the rates of illumination change in columns are the same, we

vertically stretch each column about its centre by shifting the two ends of a column. Before alignment, we derive another fusion image $\mathcal{W}_{cf}$ from the centre-aligned strip $\mathcal{W}_c$ repeating the same procedure of centre alignment. We compute the normalised shifts (the green lines in Fig. 6-3) of the upper-ends $\mathbf{B}_s$ as follows:

$$\mathbf{B}_s = \frac{1}{2}\left(1 - \frac{\sigma_a}{\mu(\sigma_a)}\right) \tag{6.3}$$

where σ_a is a vector of the column-wise standard deviations of $\mathcal{W}_{cf}$. The shifts at the bottom-ends have the same magnitude but in the opposite direction. Although our previous gradient-assisted sampling already offers a clean strip, minor alignment errors can still noticeable from the intensity distance images . These extra alignment steps alleviate these (e.g. Fig. 6-1(f) and Fig. 6-1(h)).

6.2.3 Estimation of Shadow Scale and Relighting

Using the penumbra strip, we can now derive a single unified intensity sample of each valid sub-group for estimation of penumbra illumination as opposed to performing computational-costly pixel-wise optimisation, e.g.(Liu and Gleicher, 2008; Su and Chen, 2010). Each sub-group's unified sample of intensity change G can be computed as the mean of its columns in the aligned penumbra strip such that white noise is cancelled. Unlike previous work, e.g. (Guo et al., 2012; Liu and Gleicher, 2008), assuming a constrained model of illumination change, our illumination model is adaptive to G. We fit G using a piecewise cubic Hermite polynomial (Fritsch and Carlson, 1980) which preserves the shape of the data and its monotonicity. Our illumination model is variable for different penumbra profiles as we cluster and re-scale the intensity samples in the previous steps. For each RGB channel, the normalised estimated scales are computed by dividing each estimated curve by its maximum. As the sampling sites have previously been amended during alignment of the strip, we update the two ends of each sampling line as follows:

$$\begin{bmatrix}\mathbf{p}_s^n(i)\\ \mathbf{p}_e^n(i)\end{bmatrix} = \begin{bmatrix}\mathbf{p}_s^o(i)\\ \mathbf{p}_e^o(i)\end{bmatrix} + (\mathbf{C}(i) + \begin{bmatrix}-1\\ 1\end{bmatrix}\mathbf{B}_s(i))\mathbf{V}_r(i) \tag{6.4}$$

where i is the sampling line index, $\mathbf{p}_s^o(i)$ and $\mathbf{p}_e^o(i)$ are the previous start and end points of a sampling line, $\mathbf{p}_s^n(i)$ and $\mathbf{p}_e^n(i)$ are the updated points, $\mathbf{V}_r(i)$ is the vector of boundary normal, $\mathbf{C}(i)$ and $\mathbf{B}_s(i)$ are the normalised centre and column width shifts. According to the updated sampling lines, we distribute the unified scales back to the their original positions in the image (e.g. Fig. 6-1(j)). To obtain a dense scale field (e.g. Fig. 6-1(k)), we interpolate the sparse scales in the penumbra region by smoothly interpolating and extrapolating the scales in other regions using spring-metaphor based in-painting (Bertalmio et al., 2000). The shadow-free image can be obtained by inverse scaling according to Eq. 2.4.

6.3 Evaluation

In this section, we evaluate our algorithm versus our previous two shadow removal methods (Chapter 3 and Chapter 4) and the other state-of-the-art shadow removal methods. In previous work (Guo et al., 2012; Shor and Lischinski, 2008), the quality of shadow removal is measured by directly using the per-pixel error between the shadow removal result and shadow-free ground truth. Therefore, the measurement used in Chapter 4 is adopted in this evaluation because it considers the size of the shadow, and the fact that some shadows may be darker than others. We adopt the same error ratio introduced in Eq. 4.7 as our quality measurement:

$$\mathbf{E}_r = \mathbf{E}_n / \mathbf{E}_o \tag{6.5}$$

where $\mathbf{E}_n$ is the error between the ground truth (no shadow) and shadow removal result, and $\mathbf{E}_o$ is the error between the ground truth (no shadow) and the original shadow image. This normalised measure better reflects removal improvements towards the ground truth independent of original shadow intensity and size. We assess $\mathbf{E}_n$ and $\mathbf{E}_o$ using Root-Mean-Square-Error (RMSE) of RGB intensity. To test robustness, we also compute the standard deviation for each measurement.

Improving on previous work (Guo et al., 2012; Shor and Lischinski, 2008), our removal test is based on our data set of 186 cases, which contains challenging categories for soft, broken and colour shadows and shadows cast on strong textured surfaces as well as simpler shadows, plus 28 examples from (Guo et al., 2012) – resulting in 214 test cases in total. Each case is rated according to 4 attributes, which are *texture*, *brokenness*, *colourfulness* and *softness*, in 3 perceptual degrees from weak to strong which were aggregated by five users. In Tab. 6.1 and Tab. 6.2, we show combined shadow removal error results from both automatic and semi-automatic shadow removal algorithms (all 214 cases). However, as automatic algorithms can often fail in detection – leading to artefacts unfairly biasing the removal error score – we also show results in Tab. 6.3 and Tab. 6.4 where we have removed detection failures from other algorithms – leaving 177 cases in total. This second test therefore concentrates solely on the quality of removal given accurately detected shadow inputs. As E_o for the whole image is lower than E_o for the shadow area only – and E_n for both cases are similar after shadow removal – E_r for the shadow area only is thus generally lower. In both experiments, our fast shadow removal shows leading performance across all comparisons.

		Yang et al. (2012)	Guo et al. (2012)	Su and Chen (2010)	Texture Pres. (Chap. 3)	Artefact-Resis. (Chap. 4)	Fast
Tex.	1	1.85 (1.48)	0.53 (0.50)	0.35 (0.24)	0.32 (0.19)	0.34 (0.20)	**0.26 (0.15)**
	2	1.23 (0.74)	0.59 (1.09)	0.36 (0.16)	0.38 (0.33)	0.35 (0.21)	**0.26 (0.11)**
	3	1.75 (0.89)	0.71 (0.60)	0.56 (0.29)	0.70 (0.42)	0.64 (0.60)	**0.50 (0.38)**
	M	1.61 (1.04)	0.61 (0.73)	0.42 (0.23)	0.47 (0.31)	0.44 (0.34)	**0.21 (0.22)**
Sof.	1	1.27 (0.75)	0.52 (1.08)	0.33 (0.21)	0.33 (0.31)	0.34 (0.22)	**0.23 (0.09)**
	2	1.89 (1.60)	0.70 (0.36)	0.42 (0.11)	0.44 (0.21)	0.37 (0.17)	**0.35 (0.15)**
	3	1.86 (1.12)	1.09 (0.75)	0.72 (0.21)	0.76 (0.35)	0.85 (0.58)	**0.63 (0.25)**
	M	1.67 (1.16)	0.77 (0.73)	0.49 (0.18)	0.51 (0.29)	0.52 (0.32)	**0.40 (0.17)**
Bro.	1	1.47 (1.13)	0.59 (0.98)	0.34 (0.15)	0.36 (0.29)	0.35 (0.20)	**0.26 (0.13)**
	2	1.07 (0.17)	0.42 (0.29)	0.49 (0.38)	0.44 (0.25)	0.36 (0.23)	**0.28 (0.11)**
	3	2.32 (0.96)	1.42 (1.06)	0.85 (0.25)	0.98 (0.31)	1.08 (0.66)	**0.72 (0.29)**
	M	1.62 (0.75)	0.81 (0.78)	0.56 (0.26)	0.59 (0.29)	0.59 (0.36)	**0.42 (0.17)**
Col.	1	1.20 (0.48)	0.48 (0.64)	0.34 (0.18)	0.32 (0.18)	0.33 (0.17)	**0.25 (0.10)**
	2	4.19 (2.02)	1.67 (2.29)	0.52 (0.24)	0.83 (0.67)	0.61 (0.32)	**0.48 (0.16)**
	3	2.95 (1.95)	1.20 (0.99)	0.63 (0.49)	1.10 (0.68)	1.11 (1.16)	**0.59 (0.27)**
	M	2.78 (1.48)	1.12 (1.31)	0.50 (0.30)	0.75 (0.51)	0.68 (0.55)	**0.44 (0.18)**
Other		1.72 (1.26)	0.72 (0.89)	0.47 (0.30)	0.57 (0.46)	0.56 (0.60)	**0.40 (0.29)**

Table 6.1: *Shadow removal errors for all (214) test cases. The error score in this table is computed for all pixels in the image. For each score of each attribute, the images with other predominant attributes (degree = 3) are not used. Hence, test cases have a strong single bias towards one of the attributes. "Other" refers to a set of shadow cases showing no markedly predominant attributes (degree = 1). "M" refers to the average score for each category. Standard deviations are shown in brackets. Method (Guo et al., 2012) is trained using a large shadow detection data set from (Zhu et al., 2010). The best scores are made bold.*

		Yang et al. (2012)	Guo et al. (2012)	Su and Chen (2010)	Texture Pres. (Chap. 3)	Artefact-Resis. (Chap. 4)	Fast
Tex.	1	1.81 (2.65)	0.42 (0.57)	0.17 (0.34)	0.16 (0.18)	0.17 (0.17)	**0.10 (0.10)**
	2	0.82 (0.89)	0.47 (1.15)	0.21 (0.27)	0.27 (0.35)	0.22 (0.21)	**0.13 (0.10)**
	3	1.37 (1.64)	0.64 (1.03)	0.44 (0.49)	0.65 (0.55)	0.52 (0.69)	**0.39 (0.45)**
	M	1.49 (1.36)	0.51 (0.92)	0.27 (0.37)	0.36 (0.36)	0.30 (0.36)	**0.34 (0.21)**
Sof.	1	0.80 (0.87)	0.39 (1.13)	0.18 (0.33)	0.21 (0.32)	0.21 (0.22)	**0.10 (0.09)**
	2	2.09 (2.85)	0.64 (0.43)	0.24 (0.14)	0.29 (0.25)	0.19 (0.16)	**0.16 (0.11)**
	3	1.86 (1.27)	1.01 (0.97)	0.69 (0.54)	0.71 (0.53)	0.74 (0.66)	**0.44 (0.25)**
	M	1.58 (1.66)	0.68 (0.84)	0.37 (0.34)	0.40 (0.37)	0.38 (0.34)	**0.23 (0.15)**
Bro.	1	1.19 (1.83)	0.48 (1.04)	0.18 (0.24)	0.23 (0.31)	0.21 (0.20)	**0.12 (0.10)**
	2	0.66 (0.19)	0.27 (0.35)	0.38 (0.58)	0.29 (0.28)	0.20 (0.25)	**0.11 (0.08)**
	3	2.56 (1.48)	1.55 (1.84)	0.86 (0.65)	1.05 (0.50)	1.00 (0.76)	**0.53 (0.28)**
	M	1.47 (1.17)	0.76 (1.08)	0.47 (0.49)	0.52 (0.36)	0.47 (0.40)	**0.25 (0.15)**
Col.	1	0.73 (0.45)	0.36 (0.78)	0.18 (0.23)	0.19 (0.20)	0.19 (0.18)	**0.11 (0.08)**
	2	6.07 (3.28)	1.56 (2.07)	0.45 (0.65)	0.67 (0.73)	0.42 (0.30)	**0.27 (0.17)**
	3	3.53 (3.59)	1.34 (2.33)	0.54 (0.84)	1.20 (1.18)	1.25 (1.69)	**0.49 (0.39)**
	M	3.44 (2.44)	1.09 (1.73)	0.39 (0.57)	0.69 (0.70)	0.62 (0.72)	**0.29 (0.21)**
Other		1.56 (2.05)	0.65 (1.26)	0.35 (0.50)	0.48 (0.64)	0.46 (0.78)	**0.26 (0.32)**

Table 6.2: *Shadow removal errors for all (214) test cases. The error score in this table is computed for shadow pixels in the image. For each score of each attribute, the images with other predominant attributes (degree = 3) are not used. Hence, test cases have a strong single bias towards one of the attributes. "Other" refers to a set of shadow cases showing no markedly predominant attributes (degree = 1). "M" refers to the average score for each category. Standard deviations are shown in brackets. Method (Guo et al., 2012) is trained using a large shadow detection data set from (Zhu et al., 2010). The best scores are made bold.*

		Yang et al. (2012)	Guo et al. (2012)	Su and Chen (2010)	Texture Pres. (Chap. 3)	Artefact-Resis. (Chap. 4)	Fast
Tex.	1	1.26 (0.43)	0.32 (0.25)	0.33 (0.26)	0.26 (0.10)	0.30 (0.18)	**0.23 (0.11)**
	2	1.16 (0.46)	0.43 (0.71)	0.34 (0.14)	0.35 (0.27)	0.34 (0.18)	**0.25 (0.09)**
	3	1.71 (0.84)	0.64 (0.50)	0.54 (0.26)	0.68 (0.42)	0.62 (0.63)	**0.50 (0.40)**
	M	1.38 (0.58)	0.46 (0.49)	0.40 (0.22)	0.43 (0.26)	0.42 (0.33)	**0.33 (0.20)**
Sof.	1	1.19 (0.46)	0.37 (0.65)	0.33 (0.19)	0.31 (0.24)	0.33 (0.19)	**0.23 (0.08)**
	2	1.20 (0.42)	0.51 (0.35)	0.38 (0.12)	0.39 (0.17)	0.30 (0.12)	**0.30 (0.13)**
	3	1.65 (0.83)	0.96 (0.49)	0.71 (0.21)	0.79 (0.40)	0.98 (0.65)	**0.66 (0.30)**
	M	1.35 (0.57)	0.62 (0.50)	0.47 (0.17)	0.49 (0.27)	0.54 (0.32)	**0.40 (0.17)**
Bro.	1	1.21 (0.47)	0.41 (0.63)	0.32 (0.13)	0.32 (0.24)	0.33 (0.17)	**0.24 (0.10)**
	2	1.03 (0.19)	0.29 (0.11)	0.53 (0.43)	0.34 (0.17)	0.36 (0.26)	**0.28 (0.10)**
	3	2.07 (0.67)	1.10 (0.53)	0.78 (0.18)	0.98 (0.35)	1.12 (0.72)	**0.73 (0.32)**
	M	1.43 (0.44)	0.60 (0.43)	0.54 (0.25)	0.55 (0.25)	0.60 (0.38)	**0.41 (0.17)**
Col.	1	1.19 (0.45)	0.40 (0.62)	0.33 (0.18)	0.30 (0.14)	0.33 (0.18)	**0.24 (0.09)**
	2	1.47 (0.32)	0.45 (0.00)	0.40 (0.14)	1.32 (0.82)	0.44 (0.09)	**0.38 (0.10)**
	3	3.03 (1.95)	1.21 (1.01)	0.63 (0.50)	1.12 (0.69)	1.14 (1.18)	**0.59 (0.27)**
	M	1.90 (0.91)	0.69 (0.54)	0.45 (0.28)	0.91 (0.55)	0.63 (0.48)	**0.40 (0.16)**
Other		1.58 (1.07)	0.59 (0.68)	0.45 (0.30)	0.54 (0.47)	0.56 (0.64)	**0.39 (0.30)**

Table 6.3: *Shadow removal errors for 177 test cases without detection failures. The error score in this table is computed for all pixels in the image. For each score of each attribute, the images with other predominant attributes (degree = 3) are not used. Hence, test cases have a strong single bias towards one of the attributes. "Other" refers to a set of shadow cases showing no markedly predominant attributes (degree = 1). "M" refers to the average score for each category. Standard deviations are shown in brackets. Method (Guo et al., 2012) is trained using a large shadow detection data set from (Zhu et al., 2010). The best scores are made bold.*

		Yang et al. (2012)	Guo et al. (2012)	Su and Chen (2010)	Texture Pres. (Chap. 3)	Artefact-Resis. (Chap. 4)	Fast
Tex.	1	0.70 (0.33)	0.18 (0.29)	0.17 (0.39)	0.11 (0.08)	0.15 (0.17)	**0.08 (0.05)**
	2	0.73 (0.49)	0.31 (0.88)	0.17 (0.11)	0.23 (0.31)	0.21 (0.20)	**0.12 (0.08)**
	3	1.43 (1.18)	0.48 (0.56)	0.39 (0.33)	0.65 (0.57)	0.52 (0.72)	**0.40 (0.47)**
	M	0.95 (0.67)	0.32 (0.58)	0.24 (0.28)	0.33 (0.32)	0.29 (0.37)	**0.20 (0.20)**
Sof.	1	0.70 (0.45)	0.24 (0.81)	0.16 (0.25)	0.19 (0.28)	0.20 (0.20)	**0.10 (0.07)**
	2	0.83 (0.41)	0.42 (0.42)	0.21 (0.11)	0.23 (0.19)	0.16 (0.17)	**0.13 (0.07)**
	3	1.60 (1.02)	0.80 (0.65)	0.55 (0.25)	0.76 (0.61)	0.88 (0.72)	**0.47 (0.29)**
	M	1.04 (0.63)	0.49 (0.62)	0.31 (0.20)	0.39 (0.36)	0.41 (0.36)	**0.23 (0.15)**
Bro.	1	0.73 (0.46)	0.29 (0.79)	0.14 (0.10)	0.19 (0.27)	0.19 (0.19)	**0.10 (0.08)**
	2	0.64 (0.20)	0.11 (0.06)	0.45 (0.67)	0.22 (0.19)	0.21 (0.29)	**0.10 (0.07)**
	3	2.18 (0.92)	0.93 (0.72)	0.66 (0.21)	1.06 (0.55)	1.04 (0.80)	**0.53 (0.31)**
	M	1.18 (0.53)	0.44 (0.52)	0.41 (0.33)	0.49 (0.34)	0.48 (0.43)	**0.25 (0.15)**
Col.	1	0.71 (0.43)	0.27 (0.76)	0.17 (0.23)	0.17 (0.15)	0.19 (0.19)	**0.10 (0.07)**
	2	1.33 (0.64)	0.34 (0.11)	0.21 (0.09)	1.37 (0.95)	0.55 (0.23)	**0.17 (0.05)**
	3	3.63 (3.64)	1.35 (2.38)	0.54 (0.86)	1.22 (1.20)	1.29 (1.72)	**0.49 (0.40)**
	M	1.89 (1.57)	0.66 (1.09)	0.31 (0.40)	0.92 (0.77)	0.68 (0.71)	**0.26 (0.17)**
Other		1.31 (1.74)	0.48 (1.10)	0.30 (0.42)	0.47 (0.67)	0.47 (0.84)	**0.26 (0.34)**

Table 6.4: *Shadow removal errors for 177 test cases without detection failures. The error score in this table is computed for shadow pixels in the image. For each score of each attribute, the images with other predominant attributes (degree = 3) are not used. Hence, test cases have a strong single bias towards one of the attributes. "Other" refers to a set of shadow cases showing no markedly predominant attributes (degree = 1). "M" refers to the average score for each category. Standard deviations are shown in brackets. Method (Guo et al., 2012) is trained using a large shadow detection data set from (Zhu et al., 2010). The best scores are made bold.*

Tab. 6.5 shows some typical visual results of shadow removal on various scenarios from our data set. Appendix A shows a wide range of other removal results with higher resolution images. We therefore encourage readers to examine these figures as compelling evidence of the strength of our approach. As is the case with all current shadow removal methods, our method has most difficulty in extreme cases, e.g. Tab. 6.6, where shadows are highly broken, colourful, or soft. Compared with our previous artefact-resistant shadow removal method in Chapter 4, this fast shadow removal method shows better shadow removal quality for highly broken shadows and shadows with multi-colour illumination. An example of the improved highly broken shadow removal is found in row 6 of Fig. 6.5. An example of the improved multi-colour illumination shadow removal is found in column 4 of Fig. 6.6. Compared with the failure result in Fig. 4-9, this example shows more visually satisfying shadow removal without highly visible artefacts. However, its shadow boundary recovery is still not smooth and it is therefore still considered as a failure. Fig. 6.6 also shows some other failure examples in which the shadows are highly broken (column 3), soft (column 2), coloured (column 5 and 6) and cast on textured surfaces (column 1).

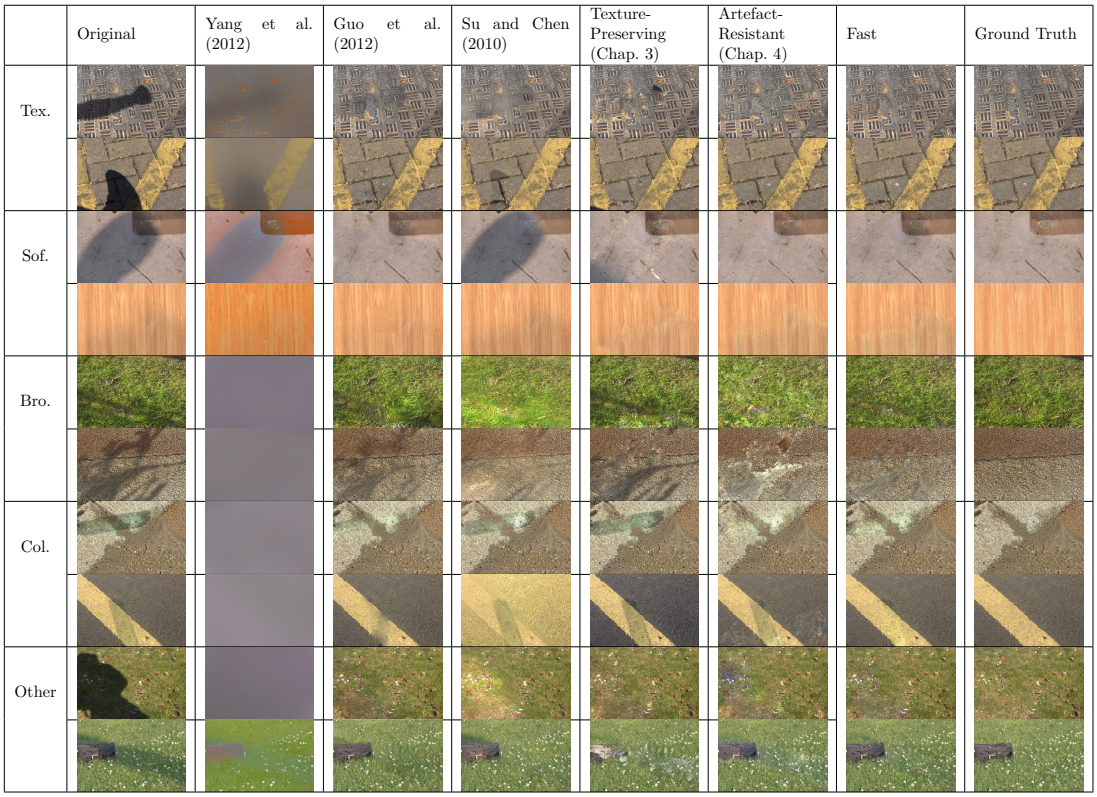


Table 6.5: *Visual comparisons using images from our data set. The table shows our results given test cases with strong degrees of the corresponding attribute except for "Other", which refers to cases where there is no predominantly strong attribute.*

Orig.						
Outp.						

Table 6.6: *Failure cases – where all the leading methods we tested in this chapter create highly visible image artefacts and fail to remove the shadow.*

6.4 Summary

We have presented an interactive method for fast shadow removal. Compared with our previous two shadow removal methods and the other previous state-of-the-art approaches, our method balances the complexity of user input with robust and fast shadow removal performance. Based on our quantitatively-verified ground truth data set, we have evaluated our method against several state-of-the-art methods using a thorough quantitative test and shown leading state-of-the-art performance. In the following Chapter 7, the concept of multi-scale processing is introduced. Without losing the advantage of fast processing speed, this multi-scale shadow removal further improves the shadow boundary recovery smoothness and robustness of shadow removal quality and a more thorough analysis of the result is performed.

Publications

The following publication is related to this chapter:
Gong, H. and Cosker, D. (2014c). Interactive shadow removal and ground truth for variable scene categories. In *British Machine Vision Conference (BMVC)*. Best Student Paper

Chapter 7

Multi-Scale Shadow Removal

Overview

In the previous Chapter 6, we have introduced a fast shadow removal method with a robust user-interaction. In this chapter, a multi-scale shadow removal method is proposed to improve shadow removal quality for a range of difficult cases: such as highly-textured and coloured shadows. This method is extended from our previous fast shadow removal presented in Chapter 6. Compared with the previous method, this approach provides finer sample alignment and adopts the concept of multi-scale processing to improve the shadow removal results in detail. As for shadow detection, the same method used in Chapter 6 is adopted. After detection, our shadow removal is performed by registering the penumbra to a normalised frame which allows efficient multi-scale estimation of non-uniform shadow illumination changes, resulting in accurate and robust removal. Based on our first validated and multi-scene category ground truth for shadow removal algorithms, the most thorough comparison of state-of-the-art shadow removal methods to date is performed. The result shows our proposed new algorithm to outperform the state-of-the-art methods across several measures and shadow category.

7.1 Introduction

In this chapter, a multi-scale processing strategy is introduced to improve our previous interactive and high-quality method for fast shadow removal. The same user interaction for shadow detection presented in Chapter 6 is adopted in this chapter. Given detection, reliable multi-scale shadow removal is delivered – verified with thorough quantitative tests for different types of shadow comparing to previous state-of-the-art approaches. Compared with the evaluation in Chapter 6, analysis of shadow attributes and performance comparison are evaluated additionally in this chapter. The evaluation is based on our large high-quality and multi-scene-category ground-truth data set for the evaluation of shadow removal (Chapter 5). The approach presented represents

what we believe to be a state-of-the-art method for shadow removal, with the most robust evaluation of such methods to date across a range of difficult shadow cases.

In this chapter, the shadow image is considered as a Hadamard product of a shadow scale layer and a shadow-free image as described in Eq. 2.4.

7.1.1 Contributions

Given the previous chapter and our review of previous work, three main contributions in this chapter are proposed:
1) Multi-scale shadow removal: Based on our previous fast penumbra unwrapping (Chapter 6), a multi-scale smoothing is introduced to derive sparse shadow scales across the penumbra. This allows robust and efficient estimation of illumination changes without requiring prior training and any assumed illumination change models. This method is simple and fast yet offers state-of-the-art shadow removal quality.
2) Robust colour correction: Post-processing effects may cause inconsistency in shadow corrected areas compared with the lit areas both in tone and contrast. Unlike the previous single scale colour correction in Chapter 4, a robust multi-scale colour correction is proposed to amend these artefacts at different scales.
3) Comprehensive evaluation: A more comprehensive evaluation is performed to compare all the shadow removal methods proposed in this thesis and the other state-of-the-art methods. Besides the detailed evaluation in the previous Chapter 6, analysis of shadow attributes and performance comparison are evaluated additionally in this chapter.

To summarise, we believe these contributions to be important to this area of research due to their significant improvements over the state-of-the-art in shadow removal in a wide range of repeatable tests.

7.2 Multi-Scale Shadow Removal

In this section, our algorithm is first explained in brief. Technical details for each of its components are then expanded upon in following sections. Our algorithm consists of four steps (see Fig. 7-1):
1) Pre-processing (§7.2.1) An initial shadow mask (Fig. 7-1(b)) is detected using a KNN classifier trained from data from two rough user inputs (e.g. Fig. 7-1(a)). A *fusion image*, which magnifies illumination discontinuities around shadow boundaries, is generated by fusing channels of YCrCb colour space and suppressing texture (Fig. 7-1(c)).
2) Penumbra unwrapping (§7.2.2) Based on the detected shadow mask and fusion image, pixel intensities of sampling lines are sampled perpendicular to the shadow boundary (Fig. 7-1(d)). Noisy samples are removed and remaining columns stored as

the initial penumbra strip (Fig. 7-1(e)). The initial columns' illumination changes are also aligned (Fig. 7-1(f)) by estimating their affine transforms.

3) Relighting (§7.2.3) From the penumbra strip, a multi-scale shadow scale estimation is applied to quickly and robustly estimate the illumination change along sampling lines and derive the sparse scales for all sampled sites (Fig. 7-1(g)) which are propagated to form a dense scale field (Fig. 7-1(h)). Shadows are removed by inverse scaling using this non-uniform field (Fig. 7-1(i)).

4) Colour correction (§7.2.4) Post-processing effects may cause inconsistent tone and contrast in shadow removed areas compared with the lit areas'. Without introducing additional artefacts, a multi-scale colour correction is proposed to remove these inconsistencies (Fig. 7-1(j)).

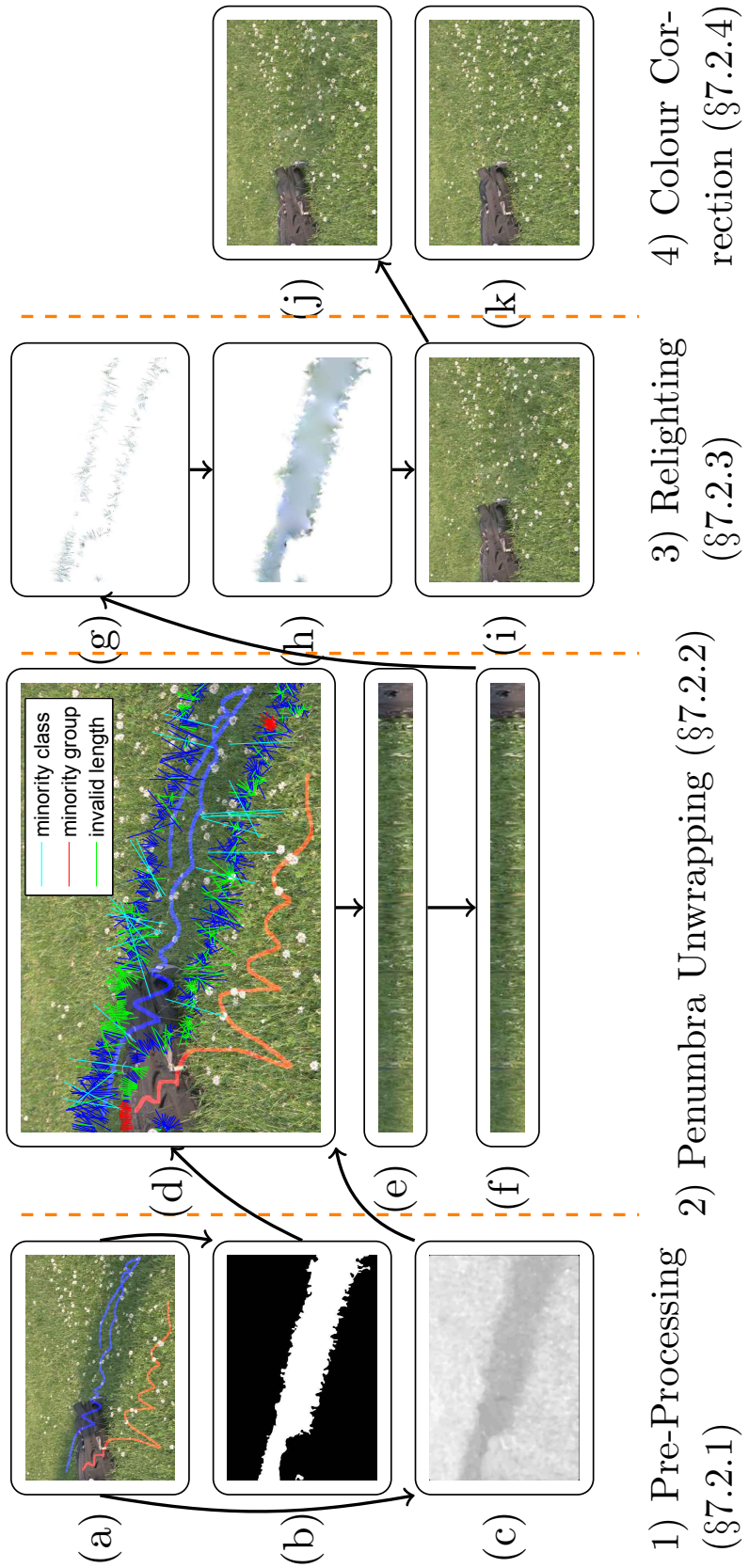


Figure 7-1: *Multi-scale shadow removal pipeline. (a) input: a shadow image and user strokes (blue for lit pixels and red for shadowed pixels); (b) detected shadow mask; (c) fusion image; (d) initial penumbra sampling (only blue lines are qualified samples); (e) initial penumbra regularisation; (f) further aligned penumbra regularisation; (g) sparse shadow scale; (h) dense shadow scale; (i) initial shadow removal result; (j) colour corrected shadow removal result; (k) ground truth.*

7.2.1 Pre-Processing

Pre-processing provides a detected shadow mask and a fusion image to assist penumbra unwrapping. Although there have been automatic methods for shadow detection, results are dependent on training data quality and variation. Instead, our method requires no prior training or learning – only two user-supplied rough inputs indicating sample lit and shadow pixels (Fig. 7-1(a)). The shadow detection algorithm used in this chapter is the same as the detection approach presented in §6.2.1.

To assist unwrapping of the penumbra, an image is derived that magnifies illumination discontinuities around the shadow boundary – also assisting penumbra location – which is called the *fusion image* (e.g. Fig. 7-1(c)). The method for deriving the fusion image is the same as the method previously described in §6.2.1.

7.2.2 Penumbra Unwrapping

A shadow boundary generally has a noisy profile with a variable penumbra width. This can lead to inaccurate estimation of shadow scales and resulting artefacts. The penumbra is therefore unwrapped into a strip and its sampled columns of illumination change are aligned (e.g. Fig. 7-1(f)). This improves the detection of outliers and allows linearisation of processing along the penumbra – leading to significant gains in efficiency and speed.

The intensity of sampling lines perpendicular to the shadow boundary (Fig. 7-1(d)) are sampled as columns of the initial penumbra strip. The length of a sampling line is determined by locating suitable start and end points guided by the fusion image $\mathcal{F}$. A bi-directional search is initialised from each boundary point that extends the sampling line towards the lit area (end point) and the shadow area (start point). The algorithm adopted in this chapter is the same as the previously described Algorithm 2. To avoid outliers, e.g. sampling lines at occlusion boundaries, invalid samples are filtered. The method for excluding outliers is based on a majority rule for log domain illumination change vectors, which is the same as the outlier exclusion algorithm previously described in §6.2.2.

As the lengths of samples are different, these are normalised by re-sizing all the samples to a unique length n_a which is the maximum length of all valid samples. The normalised samples are then concatenated as columns to form the initial penumbra strip. Although our previous adaptive sampling already provides the intensity profiles with roughly aligned illumination changes, some minor errors may still exist. Unlike the previous rough-scale alignment adopted in Chapter 6, this is resolved by a fine-scale alignment using a 1D affine transform. The parameters of an affine transform for each column are estimated by minimising the following energy function E_a:

$$E_a = \chi(\Gamma(A_s, A_k, L_o) - L_a) \tag{7.1}$$

where A_s and A_k are the linear multiplicative factors and the translation constant of 1D affine transform respectively, L_o is the scales of original column, L_a is the reference of alignment which is the average scale signal of all valid columns, Γ is a function that aligns L_o according to the estimated affine transform, χ is a function that computes sum of squares error. The minimisation is solved using a sequential quadratic programming algorithm (Nocedal and Wright, 2006).

7.2.3 Relighting

Using the aligned penumbra unwrap, a fast multi-scale shadow scale estimation is adopted for each shadow boundary. Compared with (Liu and Gleicher, 2008; Su and Chen, 2010), this shadow scale estimation is fast and adaptive, which neither requires computational-costly pixel-wise optimisation nor assumes any constrained models of illumination change, e.g. cubic curves or surface models. Compared with the previous fast shadow scale estimation in Chapter 6, this multi-scale approach provides smoother penumbra recovery.

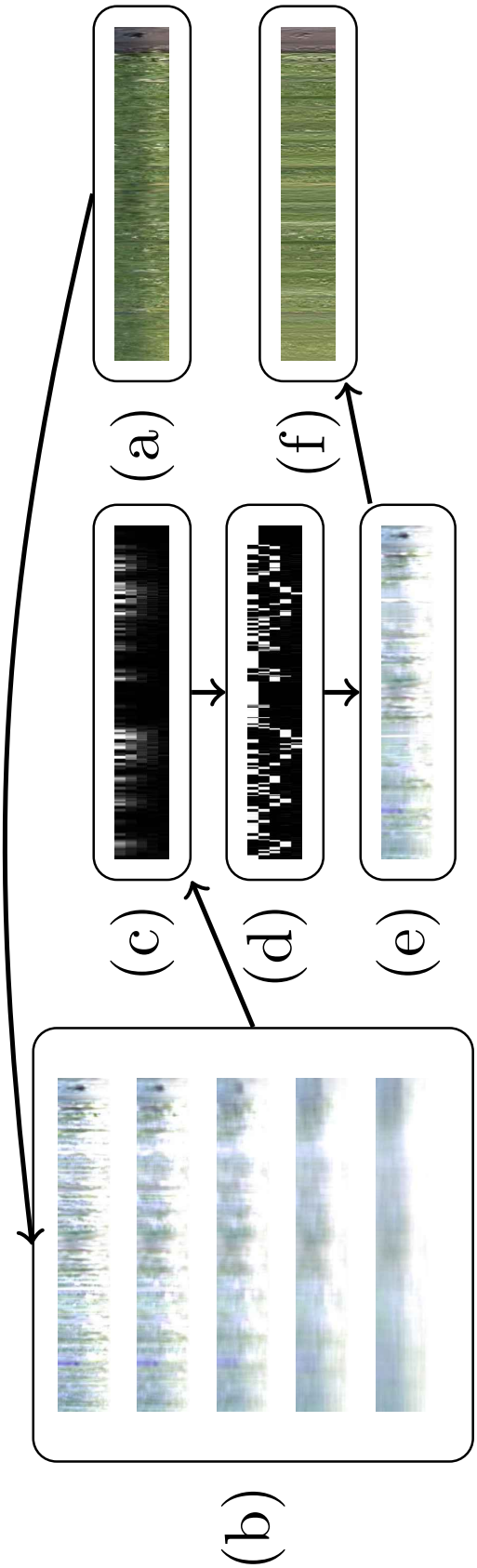


Figure 7-2: *Multi-scale shadow scale estimation pipeline. The aligned penumbra unwrap (a) is filtered using average kernels in exponentially increasing sizes to build a pyramid of shadow scales (b). The roughness of each column of each pyramid layer is measured and visualised in (c). Brighter colours indicate higher roughness. The horizontal and vertical dimensions in (c) refer to column index and layer index respectively. (d) is the visualised corresponding selections of layer index for each column (in white) after thresholding. (e) is the finale shadow scale composed by the shadow scales from the different layers of (b). (f) is the relit penumbra unwrap using (e).*

A pyramid (e.g. Fig. 7-2(b)) of horizontally filtered penumbra unwraps using 5 average kernels with different sizes are computed such that texture noise can be cancelled. The sizes of average kernels are specified as 1-by-2^n where $n \in \{2, 3, 4, 5, 6\}$. The filtered intensities of the pyramid are then converted to shadow scales. For each RGB channel layer of each pyramid layer, the estimated scales can be computed by dividing the intensities of each column by the intensity of the last element of each column (i.e. lit end). The optimum shadow scales for each column are selected from different layers of the pyramid. Column intensity with higher localness (i.e. filtered by a smaller kernel) and lower roughness are preferred. However, higher localness leads to higher roughness, so an optimum solution should balance these two properties. The roughness of intensity change $E_s(c, n)$ (visualised in Fig. 7-2(c)) is measured as follows:

$$E_s(c, n) = \int \left(\frac{\partial^2 U(r, c, n)}{\partial r^2} \right)^2 dr \tag{7.2}$$

where U is the penumbra unwrap, n is the layer index of pyramid, c and r are the column and row coordinates of the penumbra wrap respectively. The optimum scales for each column are selected using a threshold of roughness T_s which is computed as the mean value of E_s. The column of one of the layers which has the lowest roughness above T_s is selected (visualised in Fig. 7-2(d)). A shadow scale image of the penumbra unwrap (e.g. Fig. 7-2(e)) can thus be formed by picking columns from different pyramid layers according to the selected layer index of each column. As the intensity samples, i.e. columns, have previously been amended during the alignment of the unwrap, the estimated scales of each sampling are mapped back using an inverse 1D affine transform of Γ so that the estimated shadow scales are corresponding to the original unaligned intensities of the penumbra unwrap. The mapped-back shadow scales are then distributed to the their 2D positions in the image that a sparse shadow scale field is formed (e.g. Fig. 7-1(g)). To obtain a dense scale field (e.g. Fig. 7-1(h)), we propagate the sparse scales in the penumbra region by smoothly interpolating and extrapolating the scales in other regions using spring-metaphor based in-painting (Bertalmio et al., 2000). The shadow-free image can be obtained by inverse scaling according to Eq. 2.4.

7.2.4 Colour Correction

Images captured from popular imaging devices are often post-processed, e.g. gamma correction and JPEG compression, such that the linearity of photon intensity is broken. When the degree of post-processing is high, visible artefacts, e.g. differences in tone and contrast, may appear in shadow corrected areas as Eq. 2.4 does not hold. A robust multi-scale colour correction method is therefore proposed to address this issue. The difference will only be significant for images which are over post-processed. Previous work has proposed global adjustments to align the intensity characteristics of the umbra

and lit area using gradual colour transfer (§3.2.3) or gradient transfer (Liu and Gleicher, 2008). These assume that the surface around the penumbra has a similar texture and colour but may lead to significant unnatural artefacts when they are dissimilar. Unlike our previous single scale colour alignment (§4.2.4), this method aligns the spatially dependent variance of RGB intensities between the shadow and lit sides in different scales.

It is assumed that the average intensity of both sides of the shadow are accurate and that artefacts are due to the differences in intensity variance. Statistics are collected from the lit side pixels $\mathbf{P}_l$ and the umbra side pixels $\mathbf{P}_u$ both near the penumbra as the target and source of colour correction respectively. The algorithm for alignment is described in Algorithm 3. where s is a scale, β is a function that obtains the

Algorithm 3: Multi-scale colour correction

input : shadow removed image I^r, reference pixels $\mathbf{P}_l$, source pixels $\mathbf{P}_u$, shadow pixels $\mathbf{P}_s$
output: colour corrected image I_c^{ra}
$I^{ra} \leftarrow I^r$;
for $s = 1$ **to** 3 **do** scales
  $I^l \leftarrow \alpha(I^{ra}, \beta(I^{ra})/2^{s+1}, 0.2)$;
  $I^h \leftarrow I^r - I^l$;
  for $c = 1$ **to** 3 **do** RGB channels
    $r_\sigma = \varsigma(I_c^h(\mathbf{P}_l))/\varsigma(I_c^h(\mathbf{P}_u))$;
    $I_c^{ra} \leftarrow I^r$;
    $I_c^{ra}(\mathbf{P}_s) = I_c^l(\mathbf{P}_s) + r_\sigma I^h(\mathbf{P}_s)$;
  end
end

maximum image dimensional size, α is a function that bilaterally filters (Paris and Durand, 2009) the input image (first parameter) using a standard deviation of the space (second parameter) and a range Gaussian (third parameter), I^h is an image of intensity variation, where c is the channel index, ς is a function which computes the median absolute deviation.

Finally, to smooth the colour correction result, alpha blending is applied in RGB colour space according to the shadow scale as shown in Eq. 7.3.

$$I_c^f = I_c^r \circ \mathcal{S}_c^w + I_c^{ra} \circ (\mathbb{1} - \mathcal{S}_c^w) \tag{7.3}$$

where c is the channel index, $\mathcal{S}^w$ is the normalised scale field of $\mathcal{S}$, I_c^f is the final shadow-free image. An illustration of the intermediate steps of colour correction is shown in Fig 7-3.

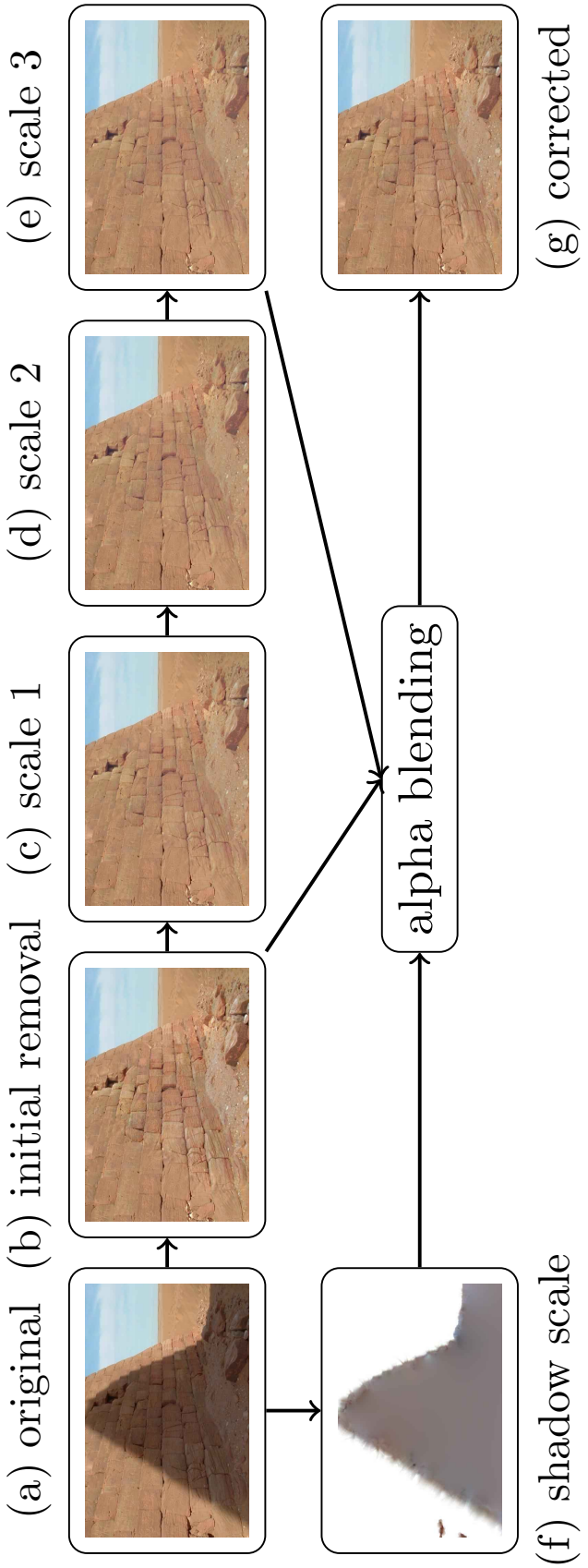


Figure 7-3: *Multi-scale colour correction pipeline. The inconsistency in the initial shadow-free image (b) is fixed in the final output (g). The multi-scale colour correction aligns the colour variance at different scales from coarse to fine.*

7.3 Evaluation

In this section, we first describe experiments that highlight our algorithms behaviour given variable user inputs. After this, our algorithm is evaluated versus other state-of-the-art shadow removal methods based on our new data set (Chapter 5).

7.3.1 Performance Stability Given Different User Inputs

Given user-supplied single pairs of strokes of lit and shadow pixel samples, our shadow detection generates stable results in different conditions (e.g. Fig. 7-4(a) and Fig. 7-4(b)). In some cases, e.g. where the surface colour is very shadow-like, the detection results can be improved by supplying more than one pair of strokes (e.g. Fig. 7-4(c) and Fig. 7-4(d)).

(a) single pair of strokes test (b) single pair of strokes test

(c) result with less strokes (d) result with more strokes

Figure 7-4: *Variable input behaviours: The top row shows two examples using single pairs of strokes. 10 examples of single strokes placed in different locations are supplied as input (red for lit and blue for shadow). The 2 grey-level images show the visualised probability of each pixel being marked in these 10 independent tests. Fewer grey pixels indicate higher stability, i.e. the image should only show black (0% probability) and white (100% probability) pixels when it is absolutely stable. The bottom row shows examples highlighting how additional strokes can improve the detection result (binary mask).*

7.3.2 Quantitative Evaluation of Shadow Removal

We consider the size of the shadow, and the fact that some shadows may be darker than others. In this chapter, we therefore adopt the error ratio previously introduced in Eq. 4.7 as our quality measurement. This normalised measure better reflects removal improvements towards the ground truth independent of original shadow intensity and size.

To test robustness, the standard deviation for each measurement is also computed. Unlike previous uncategorised test (Guo et al., 2012; Shor and Lischinski, 2008), our removal test is based on our data set of 186 cases, which contains challenging soft, broken and colour shadows and shadows cast on strong textured surfaces as well as simpler shadows, plus 28 remaining cases from (Guo et al., 2012) – resulting in 214 test cases in total. Each case is rated according to 4 attributes, which are *texture*, *brokenness*, *colourfulness* and *softness*, in 3 perceptual degrees from weak to strong which were aggregated by 5 users.

In Tab. 7.1 Tab. 7.2, Tab. 7.3, and Tab. 7.4, the combined shadow removal error results from both automatic and semi-automatic shadow removal algorithms (all 214 cases) are shown. However, as automatic algorithms can often fail in detection – leading to artefacts unfairly biasing the removal error score – we also show results in Tab. 7.3 and Tab. 7.4 where detection failures are removed from other algorithms – leaving 177 cases in total. This second test therefore concentrates solely on the quality of *removal* given accurately detected shadow inputs. In both experiments, our fast shadow removal (Chapter 6) and its extension multi-scale shadow removal in this chapter show leading performance across all scores. The quantitative results of these two method are very close. The differences between them can be examined in the extended visual comparison in Appendix A.

The scores of (Guo et al., 2012) and our texture-preserving shadow removal (Chapter 3) exhibit that their detection methods may often be unsuitable in some cases. However, they require little or no user inputs in contrast to our method. This demonstrates the careful balance of using user-assisted methods, i.e. while automatic methods are often desirable, the use of simple and robust input that gives consistent quality results may often be more suitable. Regardless of the stability of the shadow detection approaches adopted, our texture-preserving shadow removal (Chapter 3), artefact-resistant shadow removal (Chapter 4) and (Su and Chen, 2010) show similar second best overall performance although they show different degrees of capability in processing shadows given different attributes and categories.

		Guo et al. (2012)	Su and Chen (2010)	Texture-Pres. (Chap. 3)	Artefact-Resi. (Chap. 4)	Fast (Chap. 6)	Multi-Scale
Tex.	1	0.53 (0.50)	0.35 (0.24)	0.32 (0.19)	0.34 (0.20)	**0.26 (0.15)**	0.27 (0.16)
	2	0.59 (1.09)	0.36 (0.16)	0.38 (0.33)	0.35 (0.21)	**0.26 (0.11)**	**0.26 (0.11)**
	3	0.71 (0.60)	0.56 (0.29)	0.70 (0.42)	0.64 (0.60)	0.50 (0.38)	**0.49 (0.33)**
	M	0.61 (0.73)	0.42 (0.23)	0.47 (0.31)	0.44 (0.34)	0.34 (0.21)	**0.34 (0.20)**
Sof.	1	0.52 (1.08)	0.33 (0.21)	0.33 (0.31)	0.34 (0.22)	**0.23 (0.09)**	0.23 (0.10)
	2	0.70 (0.36)	0.42 (0.11)	0.44 (0.21)	0.37 (0.17)	0.35 (0.15)	**0.34 (0.15)**
	3	1.09 (0.75)	0.72 (0.21)	0.76 (0.35)	0.85 (0.58)	**0.63 (0.25)**	0.45 (0.23)
	M	0.77 (0.73)	0.49 (0.18)	0.51 (0.29)	0.52 (0.32)	**0.40 (0.17)**	0.40 (0.18)
Bro.	1	0.59 (0.98)	0.34 (0.15)	0.36 (0.29)	0.35 (0.20)	**0.26 (0.13)**	**0.26 (0.13)**
	2	0.42 (0.29)	0.49 (0.38)	0.44 (0.25)	0.36 (0.23)	**0.28 (0.11)**	0.30 (0.14)
	3	1.42 (1.06)	0.85 (0.25)	0.98 (0.31)	1.08 (0.66)	**0.72 (0.29)**	0.73 (0.28)
	M	0.81 (0.78)	0.56 (0.26)	0.59 (0.29)	0.59 (0.36)	**0.42 (0.17)**	0.43 (0.18)
Col.	1	0.48 (0.64)	0.34 (0.18)	0.32 (0.18)	0.33 (0.17)	0.25 (0.10)	**0.24 (0.11)**
	2	1.67 (2.29)	0.52 (0.24)	0.83 (0.67)	0.61 (0.32)	**0.48 (0.16)**	0.49 (0.15)
	3	1.20 (0.99)	0.63 (0.49)	1.10 (0.68)	1.11 (1.16)	0.59 (0.27)	**0.58 (0.34)**
	M	1.12 (1.31)	0.50 (0.30)	0.75 (0.51)	0.68 (0.55)	**0.44 (0.18)**	0.44 (0.20)
Other		0.72 (0.89)	0.47 (0.30)	0.57 (0.46)	0.56 (0.60)	0.40 (0.29)	**0.40 (0.28)**

Table 7.1: *Shadow removal errors for all (214) test cases. The error score in this table is computed for all pixels in the image. For each score of each attribute, the images with other predominant attributes (degree = 3) are not used. Hence, test cases have a strong single bias towards one of the attributes. "Other" refers to a set of shadow cases showing no markedly predominant attributes (degree = 1). "M" refers to the average score for each category. Standard deviations are shown in brackets. Method (Guo et al., 2012) is trained using a large shadow detection data set from (Zhu et al., 2010). The best scores are made bold.*

		Guo et al. (2012)	Su and Chen (2010)	Texture-Pres. (Chap. 3)	Artefact-Resi. (Chap. 4)	Fast (Chap. 6)	Multi-Scale
Tex.	1	0.42 (0.57)	0.17 (0.34)	0.16 (0.18)	0.17 (0.17)	**0.10 (0.10)**	**0.10 (0.10)**
	2	0.47 (1.15)	0.21 (0.27)	0.27 (0.35)	0.22 (0.21)	0.13 (0.10)	**0.13 (0.09)**
	3	0.64 (1.03)	0.44 (0.49)	0.65 (0.55)	0.52 (0.69)	0.39 (0.45)	**0.35 (0.30)**
	M	0.51 (0.92)	0.27 (0.37)	0.36 (0.36)	0.30 (0.36)	0.21 (0.22)	**0.19 (0.16)**
Sof.	1	0.39 (1.13)	0.18 (0.33)	0.21 (0.32)	0.21 (0.22)	0.10 (0.09)	**0.10 (0.08)**
	2	0.64 (0.43)	0.24 (0.14)	0.29 (0.25)	0.19 (0.16)	**0.16 (0.11)**	**0.16 (0.11)**
	3	1.01 (0.97)	0.69 (0.54)	0.71 (0.53)	0.74 (0.66)	**0.44 (0.25)**	0.45 (0.23)
	M	0.68 (0.84)	0.37 (0.34)	0.40 (0.37)	0.38 (0.34)	**0.23 (0.15)**	0.24 (0.14)
Bro.	1	0.48 (1.04)	0.18 (0.24)	0.23 (0.31)	0.21 (0.20)	0.12 (0.10)	**0.12 (0.09)**
	2	0.27 (0.35)	0.38 (0.58)	0.29 (0.28)	0.20 (0.25)	**0.11 (0.08)**	0.14 (0.11)
	3	1.55 (1.84)	0.86 (0.65)	1.05 (0.50)	1.00 (0.76)	**0.53 (0.28)**	0.54 (0.24)
	M	0.76 (1.08)	0.47 (0.49)	0.52 (0.36)	0.47 (0.40)	**0.25 (0.15)**	0.26 (0.15)
Col.	1	0.36 (0.78)	0.18 (0.23)	0.19 (0.20)	0.19 (0.18)	**0.11 (0.08)**	**0.11 (0.08)**
	2	1.56 (2.07)	0.45 (0.65)	0.67 (0.73)	0.42 (0.30)	0.27 (0.17)	**0.26 (0.13)**
	3	1.34 (2.33)	0.54 (0.84)	1.20 (1.18)	1.25 (1.69)	0.49 (0.39)	**0.47 (0.45)**
	M	1.09 (1.73)	0.39 (0.57)	0.69 (0.70)	0.62 (0.72)	0.29 (0.21)	**0.28 (0.22)**
Other		0.65 (1.26)	0.35 (0.50)	0.48 (0.64)	0.46 (0.78)	0.26 (0.32)	**0.25 (0.28)**

Table 7.2: *Shadow removal errors for all (214) test cases. The error score in this table is computed for shadow pixels in the image. For each score of each attribute, the images with other predominant attributes (degree = 3) are not used. Hence, test cases have a strong single bias towards one of the attributes. "Other" refers to a set of shadow cases showing no markedly predominant attributes (degree = 1). "M" refers to the average score for each category. Standard deviations are shown in brackets. Method (Guo et al., 2012) is trained using a large shadow detection data set from (Zhu et al., 2010). The best scores are made bold.*

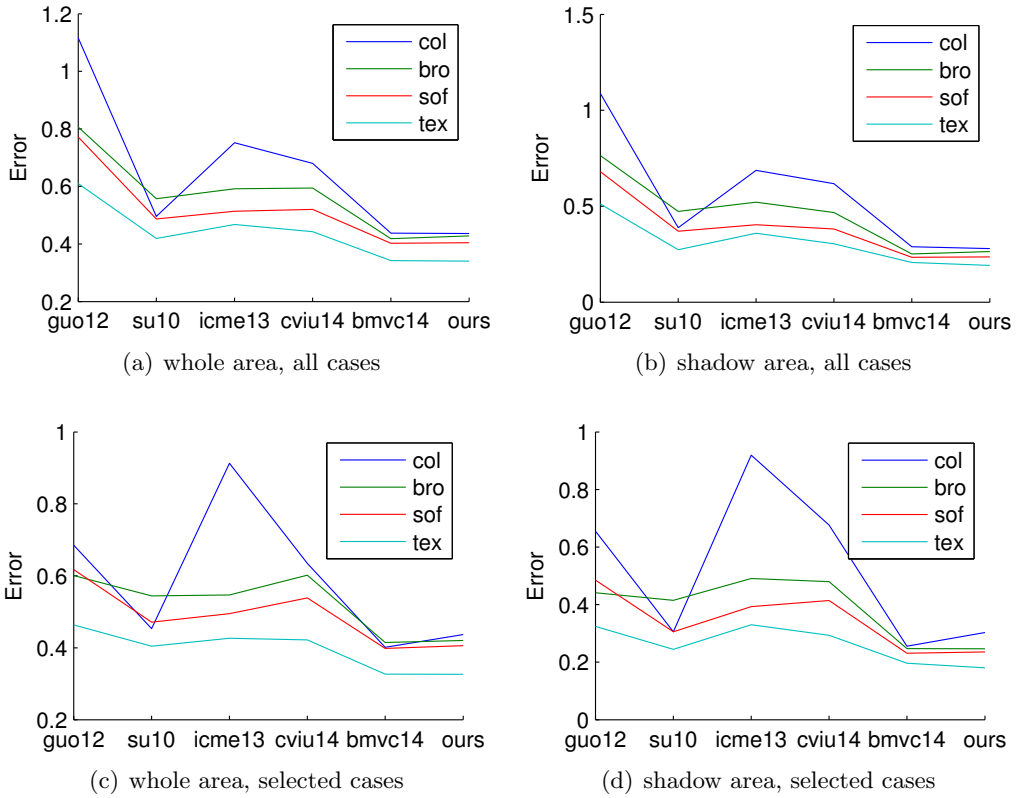
		Guo et al. (2012)	Su and Chen (2010)	Texture-Pres. (Chap. 3)	Artefact-Resi. (Chap. 4)	Fast (Chap. 6)	Multi-Scale
Tex.	1	0.32 (0.25)	0.33 (0.26)	0.26 (0.10)	0.30 (0.18)	**0.23 (0.11)**	0.23 (0.12)
	2	0.43 (0.71)	0.34 (0.14)	0.35 (0.27)	0.34 (0.18)	**0.25 (0.09)**	0.25 (0.10)
	3	0.64 (0.50)	0.54 (0.26)	0.68 (0.42)	0.62 (0.63)	0.50 (0.40)	**0.50 (0.34)**
	M	0.46 (0.49)	0.40 (0.22)	0.43 (0.26)	0.42 (0.33)	**0.33 (0.20)**	**0.33 (0.19)**
Sof.	1	0.37 (0.65)	0.33 (0.19)	0.31 (0.24)	0.33 (0.19)	**0.23 (0.08)**	0.23 (0.10)
	2	0.51 (0.35)	0.38 (0.12)	0.39 (0.17)	0.30 (0.12)	0.30 (0.13)	**0.29 (0.14)**
	3	0.96 (0.49)	0.71 (0.21)	0.79 (0.40)	0.98 (0.65)	**0.66 (0.30)**	0.69 (0.29)
	M	0.62 (0.50)	0.47 (0.17)	0.49 (0.27)	0.54 (0.32)	**0.40 (0.17)**	0.41 (0.18)
Bro.	1	0.41 (0.63)	0.32 (0.13)	0.32 (0.24)	0.33 (0.17)	**0.24 (0.10)**	0.24 (0.11)
	2	0.29 (0.11)	0.53 (0.43)	0.34 (0.17)	0.36 (0.26)	**0.28 (0.10)**	0.29 (0.14)
	3	1.10 (0.53)	0.78 (0.18)	0.98 (0.35)	1.12 (0.72)	**0.73 (0.32)**	**0.73 (0.32)**
	M	0.60 (0.43)	0.54 (0.25)	0.55 (0.25)	0.60 (0.38)	**0.41 (0.17)**	0.42 (0.19)
Col.	1	0.40 (0.62)	0.33 (0.18)	0.30 (0.14)	0.33 (0.18)	**0.24 (0.09)**	0.24 (0.10)
	2	0.45 (0.00)	0.40 (0.14)	1.32 (0.82)	0.44 (0.09)	**0.38 (0.10)**	0.48 (0.11)
	3	1.21 (1.01)	0.63 (0.50)	1.12 (0.69)	1.14 (1.18)	**0.59 (0.27)**	0.59 (0.35)
	M	0.69 (0.54)	0.45 (0.28)	0.91 (0.55)	0.63 (0.48)	**0.40 (0.16)**	0.44 (0.19)
Other		0.59 (0.68)	0.45 (0.30)	0.54 (0.47)	0.56 (0.64)	**0.39 (0.30)**	**0.39 (0.30)**

Table 7.3: *Shadow removal errors for 177 test cases without detection failures. The error score in this table is computed for all pixels in the image. For each score of each attribute, the images with other predominant attributes (degree = 3) are not used. Hence, test cases have a strong single bias towards one of the attributes. "Other" refers to a set of shadow cases showing no markedly predominant attributes (degree = 1). "M" refers to the average score for each category. Standard deviations are shown in brackets. Method (Guo et al., 2012) is trained using a large shadow detection data set from (Zhu et al., 2010). The best scores are made bold.*

		Guo et al. (2012)	Su and Chen (2010)	Texture-Pres. (Chap. 3)	Artefact-Resi. (Chap. 4)	Fast (Chap. 6)	Multi-Scale
Tex.	1	0.18 (0.29)	0.17 (0.39)	0.11 (0.08)	0.15 (0.17)	**0.08 (0.05)**	0.08 (0.06)
	2	0.31 (0.88)	0.17 (0.11)	0.23 (0.31)	0.21 (0.20)	**0.12 (0.08)**	0.12 (0.09)
	3	0.48 (0.56)	0.39 (0.33)	0.65 (0.57)	0.52 (0.72)	0.40 (0.47)	**0.34 (0.30)**
	M	0.32 (0.58)	0.24 (0.28)	0.33 (0.32)	0.29 (0.37)	0.20 (0.20)	**0.18 (0.15)**
Sof.	1	0.24 (0.81)	0.16 (0.25)	0.19 (0.28)	0.20 (0.20)	**0.10 (0.07)**	0.10 (0.08)
	2	0.42 (0.42)	0.21 (0.11)	0.23 (0.19)	0.16 (0.17)	**0.13 (0.07)**	0.13 (0.09)
	3	0.80 (0.65)	0.55 (0.25)	0.76 (0.61)	0.88 (0.72)	**0.47 (0.29)**	0.48 (0.24)
	M	0.49 (0.62)	0.31 (0.20)	0.39 (0.36)	0.41 (0.36)	**0.23 (0.15)**	0.24 (0.14)
Bro.	1	0.29 (0.79)	0.14 (0.10)	0.19 (0.27)	0.19 (0.19)	**0.10 (0.08)**	0.11 (0.08)
	2	0.11 (0.06)	0.45 (0.67)	0.22 (0.19)	0.21 (0.29)	**0.10 (0.07)**	0.12 (0.09)
	3	0.93 (0.72)	0.66 (0.21)	1.06 (0.55)	1.04 (0.80)	0.53 (0.31)	**0.51 (0.25)**
	M	0.44 (0.52)	0.41 (0.33)	0.49 (0.34)	0.48 (0.43)	0.25 (0.15)	textbf0.25 (0.14)
Col.	1	0.27 (0.76)	0.17 (0.23)	0.17 (0.15)	0.19 (0.19)	**0.10 (0.07)**	0.10 (0.08)
	2	0.34 (0.11)	0.21 (0.09)	1.37 (0.95)	0.55 (0.23)	**0.17 (0.05)**	0.33 (0.03)
	3	1.35 (2.38)	0.54 (0.86)	1.22 (1.20)	1.29 (1.72)	0.49 (0.40)	**0.48 (0.46)**
	M	0.66 (1.09)	0.31 (0.40)	0.92 (0.77)	0.68 (0.71)	**0.26 (0.17)**	0.30 (0.19)
Other		0.48 (1.10)	0.30 (0.42)	0.47 (0.67)	0.47 (0.84)	0.26 (0.34)	**0.25 (0.29)**

Table 7.4: *Shadow removal errors for 177 test cases without detection failures. The error score in this table is computed for shadow pixels in the image. For each score of each attribute, the images with other predominant attributes (degree = 3) are not used. Hence, test cases have a strong single bias towards one of the attributes. "Other" refers to a set of shadow cases showing no markedly predominant attributes (degree = 1). "M" refers to the average score for each category. Standard deviations are shown in brackets. Method (Guo et al., 2012) is trained using a large shadow detection data set from (Zhu et al., 2010). The best scores are made bold.*

7.3.3 Analysis of Shadow Categories and Attributes



(a) whole area, all cases

(b) shadow area, all cases

(c) whole area, selected cases

(d) shadow area, selected cases

Figure 7-5: *Parallel coordinate charts of the quantitative results in Tab. 7.1 and Tab. 7.2. The ticks guo12, su10, icme13, cviu14, bmvc14, ours refer to (Guo et al., 2012), (Su and Chen, 2010), texture-preserving shadow removal (Chapter 3), artefact-resistant shadow removal (Chapter 4), fast shadow removal (Chapter 6), and multi-scale shadow removal (this chapter) accordingly. The selected cases refer to those cases without significant detection failures for all tested methods. The scores presented here are the average scores of all three degrees for each attribute.*

To investigate the affects of different shadow categories and attributes, the quantitative results in Tab. 7.1 Tab. 7.2, Tab. 7.3, and Tab. 7.4 are summarised by visualising the result using the parallel coordinate charts in Fig. 7-5. Such a visualisation is insightful as strong performance of one method could direct practitioners to favour one algorithm over another in some problem cases. Overall, coloured shadows are shown to be significantly the most difficult shadows to remove and shadows cast on high texture the easiest challenge. Broken shadows are slightly more difficult to process than soft shadows, although both of them are in the range of medium difficulty. (Guo et al., 2012), our texture-preserving shadow removal (Chapter 3) and our artefact-resistant shadow removal (Chapter 4) show relatively significant disadvantages in processing colour shadows, while all methods, expect for our fast shadow removal (Chapter 6) and our multi-scale shadow removal (this chapter), demonstrate obvious difficulty in pro-

cessing broken shadows. The trend of the other methods and attributes are otherwise similar. In our tests, our fast shadow removal (Chapter 6) and multi-scale shadow removal (this chapter) overall demonstrate the best performance for all types of shadows analysed.

7.3.4 Visual Comparison

Fig. 7-6 shows some typical visual results of shadow removal on various scenarios from our data set. Appendix A shows a wide range of other removal results with higher resolution images. We therefore encourage readers to examine these figures as compelling evidence of the strength of our approach. Compared with our previous fast shadow removal (Chapter 6), this multi-scale method has improved the colour shadow removal quality and achieved smoother shadow removal. For example, the previous failure cases of colour shadow removal in column 4, 5 and 6 of Fig. 6.6 are better processed. Although these colour shadow removal results are still with some light traces of colour shadows, the illumination processing of this method is very smooth. As is the case with all current shadow removal methods, this algorithm has most difficulty in some extreme cases, e.g. Fig. 7-7, where shadows are highly-broken, colourful, or soft.

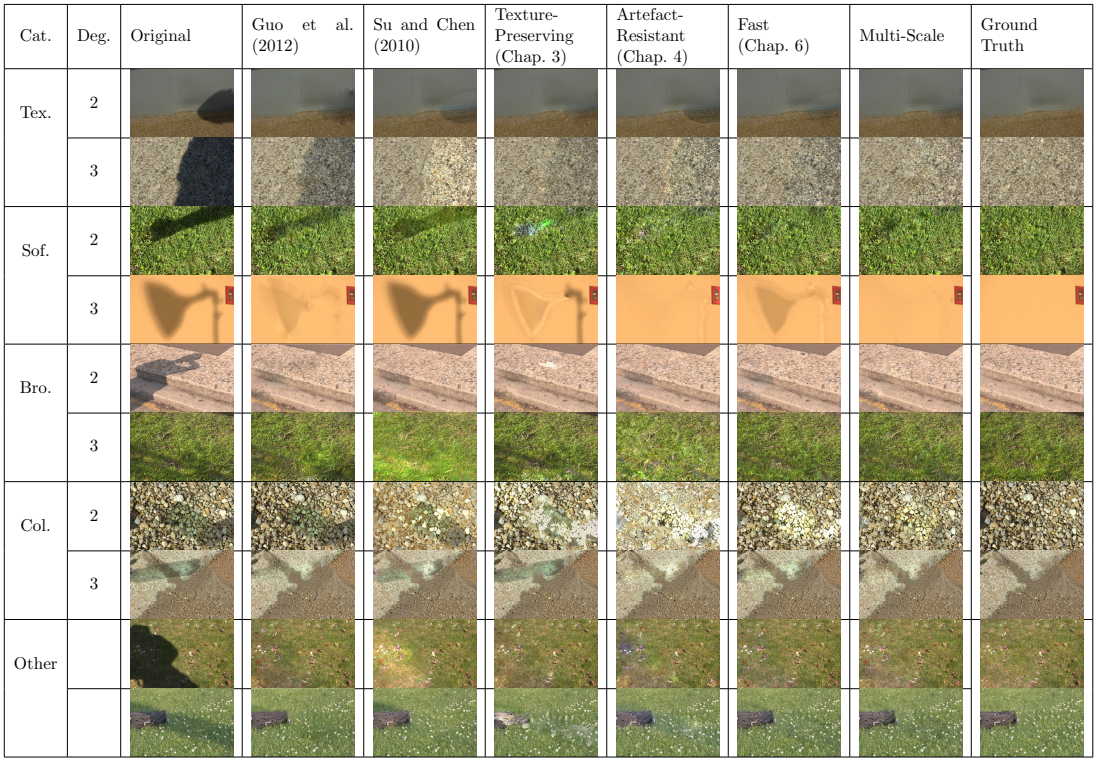


Figure 7-6: *Visual comparisons using images from our data set. The table shows our results given test cases with stronger degrees of the corresponding attribute except for "Other", which refers to cases where there is no predominantly strong attribute. Degree 1 is not shown as 1 means almost no attribute feature.*

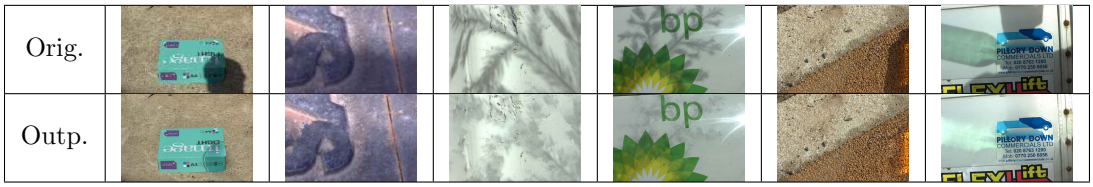


Figure 7-7: *Failure cases – where all the leading methods tested result in highly visible image artefacts and fail to remove the shadow. These highlight directions for future work in our field.*

7.3.5 Efficiency Comparison

Tab. 7.5 shows the required time for processing 0.3 mega-pixel colour images shown in Tab. 7-6 on a 3.1GHz machine. Our MATLAB implementation of fast shadow removal and its extension – multi-scale shadow removal in this chapter generally require less system processing time than the other three MATLAB implementations and one C/C++ implementation. Our artefact-resistant shadow removal (Chapter 4) processes significantly faster than our texture-preserving shadow removal (Chapter 3). As for user-interaction, our artefact-resistant shadow removal is the easiest. The same user-interaction adopted in fast shadow removal (Chapter 6) and multi-scale shadow removal (Chapter 7) also require little time for user-interaction compared with the other two user-assisted methods. Our texture-preserving shadow removal requires little time for user-interaction for simple scenes however it requires more time for the images contain more shadow regions.

		Guo et al. (2012) MATLAB (MEX)			Su and Chen (2010) C/C++			Texture-Preserving (Chap. 3) MATLAB			Artefact-Resistant (Chap. 4) MATLAB			Fast (Chap. 6) MATLAB			Multi-Scale MATLAB		
		T_u	T_s	T_t	T_u	T_s	T_t	T_u	T_s	T_t	T_u	T_s	T_t	T_u	T_s	T_t	T_u	T_s	T_t
Tex.	2	0	157	157	22	44	66	2	35	37	2	10	12	4	6	10	4	7	11
	3	0	74	74	20	44	64	2	81	83	2	12	24	4	8	12	4	10	14
Sof.	2	0	159	159	18	44	62	4	267	271	3	8	11	4	11	15	4	13	17
	3	0	23	23	109	43	152	9	31	40	2	26	28	3	15	18	3	21	24
Bro.	2	0	61	61	37	43	80	8	88	96	3	28	31	4	11	15	4	14	18
	3	0	158	158	61	51	112	10	241	251	3	139	142	8	43	51	8	54	62
Col.	2	0	160	160	30	45	75	3	385	388	3	264	267	5	42	47	5	51	56
	3	0	56	56	19	46	65	2	132	134	2	27	29	5	8	13	5	11	16
Other		0	74	74	33	42	75	4	301	305	2	17	19	4	10	14	4	17	21
		0	80	80	28	42	70	2	114	116	3	59	62	4	14	18	4	13	17

Table 7.5: *Time comparison of shadow removal: T_u, T_s, T_t refer to time (in seconds) for user interaction, time for system processing, and total time respectively. The type of implementation can affect the performance significantly and these are denoted below the corresponding methods. MEX refers to a C/C++ optimised MATLAB function.*

7.4 Summary

We have presented an interactive method for multi-scale shadow removal. Our method balances the complexity of user input with robust shadow removal performance. We have evaluated our method against our proposed methods in all the previous chapters and some state-of-the-art methods using our thorough quantitative test and shown

leading state-of-the-art performance. Besides the opportunities for improving shadow removal quality for the categorised shadows in our data set, the removal for highly-complicated shadows, such as overlapping shadows caused multiple light sources with different light colours, and shadows caused by transparent objects with complicated inner structure and colour, is still an open problem for the community. In the following Chapter 8, a novel application of interactive shadow editing, which is based on this multi-scale shadow removal method, is presented.

Publications

- Gong, H. and Cosker, D. (2014b). Interactive shadow removal and ground truth for variable scene categories. *International Journal of Computer Vision (IJCV)*. under revision

Chapter 8

Application: Interactive Shadow Editing from Single Images

Overview

We present a system for interactive shadow editing from single images which includes the manipulations of shape, distribution, sharpness and darkness of shadows according to the features of existing shadows. We first obtain a shadow-free image, shadow boundary and its registered sparse shadow scales using our multi-scale shadow removal method (Chapter 7). The modifiable features of the shadow are synthesised from the sparse shadow scales. According to the user-specified shadow-shape and its attributes, our system generates a new shadow matte and composites it into the original image, while also allowing editing of existing shadows.

8.1 Introduction

In this chapter, we propose a system for interactive shadow editing from single images which synthesises the features of existing shadow and preserves the naturalness of newly generated shadows. Our system provides friendly and flexible user-controls for defining the shape, darkness, softness, and colour of existing or new shadows. Furthermore, it does not require users to manually analyse and adjust the shadow properties in an image, as the manipulatable shadow properties are automatically synthesised by our system. Compared with the artificial shadows generated by Computer Graphics rendering, our shadow modification is based on the existing real shadows and our newly modified and added shadows can be highly consistent with the original shadows in a scene. The potential usages of this tool include but are not limited to: (1) moving existing objects and their shadows from one image to another; and (2) artistic modification of shadow.

This chapter focuses on a shadow editing framework that utilises the output information from an existing shadow extraction method. Based on our state-of-the-art shadow removal data set (Chapter 5), we have both visually and quantitatively tested our system on various shadow scenes, which includes scenes with strong texture background, broken, soft and coloured shadows. In this chapter, the shadow effects are represented as a Hadamard product of a shadow scale field and a shadow-free image as described in Eq. 2.4.

8.1.1 Contributions

Given our review of state-of-the-art approaches, we propose the following two contributions:
1) A model for shadow editing from single images We propose a model that analyses the features of an existing shadow in an image and provides parameters for users to edit the extracted shadow and add new shadows. Unlike previous work (Mohan et al., 2007), our shadow editing model is fast, highly flexible and provides many more controls of shadow properties. The model is also universal and compatible with several shadow removal methods.
2) Easy user interaction design for shadow editing An easy interface is proposed for users to freely and quickly define the shape of shadow and control various properties of shadow.

To summarise, we believe our contributions are important in this area due to our significant improvements in extracting and utilising controllable properties from existing shadow and the ease in which a user may interact with a shadow scene.

8.2 Interactive Shadow Editing Model

In this section, we overview our algorithm first in brief then expend on technical details for each step. Our algorithm consists of three steps (see Fig. 8-1):
1) Pre-processing (§8.2.1) A shadow-free image, shadow boundary information, and shadow scale attenuation profiles are first obtained. This information is converted into a field of shadow softness and a synthesised intensity attenuation profile.
2) Synthesis of shadow matte (§8.2.2) Based on the pre-processed information, a shadow model with various tunable parameters is generated. A modified shadow matte is synthesised using a distance transform.
3) Composition of shadow (§8.2.3) The modified shadow is seamlessly composited into the original shadow image.

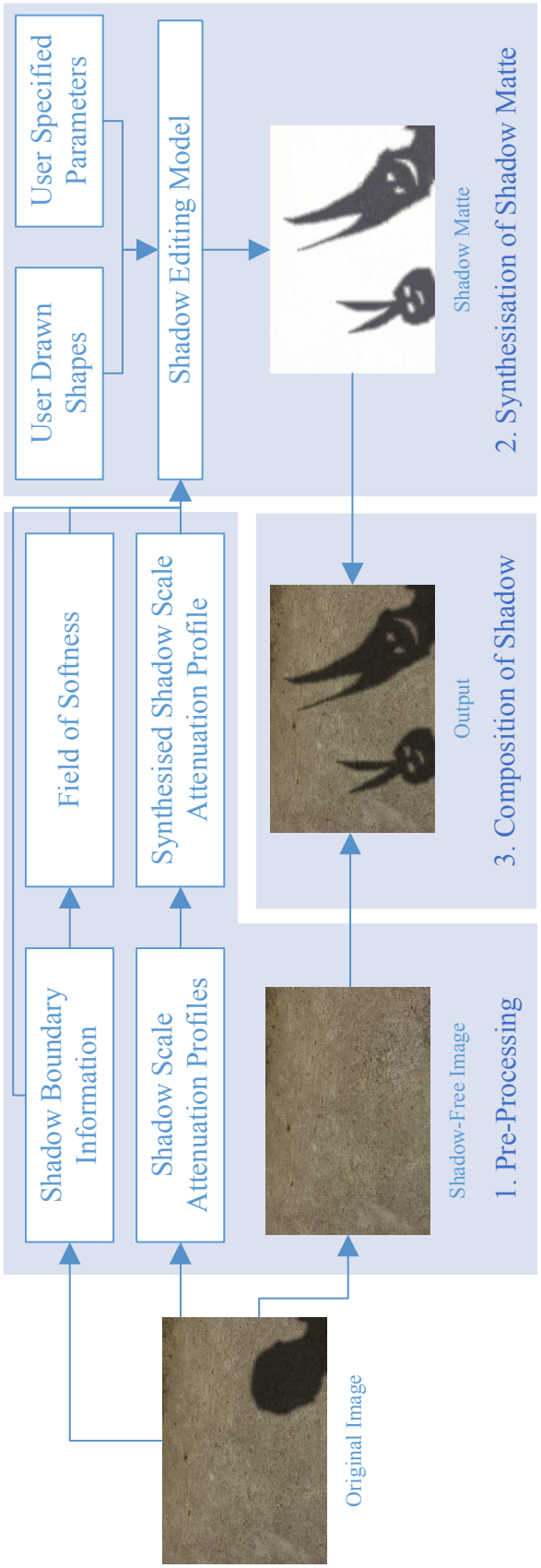


Figure 8-1: *Shadow editing pipeline. The three main blocks in this chart correspond to the three main steps of our shadow editing system.*

8.2.1 Pre-Processing

Pre-processing provides a shadow-free image, shadow boundaries, a field of shadow softness and a synthesised shadow scale attenuation profile required by our shadow editing model.

Shadow Extraction

Shadow extraction is the first step for shadow editing which provides necessary information for realistic shadow re-creation. To extract shadow from a single image, our multi-scale shadow removal method (Chapter 7) is applied. This shadow removal method gives a shadow-free image, a shadow mask, and many shadow scale attenuation profiles perpendicular to the shadow boundary. All our previous methods and some other shadow removal methods (Arbel and Hel-Or, 2011; Liu and Gleicher, 2008; Mohan et al., 2007; Su and Chen, 2010) which rely on intensity profile analysis are also compatible to our system.

Synthesisation of Shadow Scale Attenuation

Shadow creation is simplified by synthesising intensity attenuation. All extracted shadow scale profiles are re-scaled to a unique length (length of the longest profile by default). A synthesised shadow scale attenuation profile is then generated by profile-wise averaging all re-scaled shadow scale profiles. The averaging process cancels texture noise. To accelerate the computation for variable penumbra widths, a look-up table like profile is pre-computed. The synthesised shadow scale attenuation profile is evenly re-sampled to a large number (1000 in our implementation) of data sites using a piece-wise cubic Hermite polynomial (Fritsch and Carlson, 1980). Shadow scales on wide penumbra can thus be queried by finding the closest scale from the large number of data sites.

Generation of Softness Field

When the shape of shadow is changed from the original shadow image, it is unknown how soft the penumbra of the newly created parts should be. This is solved by generating a field of shadow softness from known penumbra widths of original shadow boundary points. The problem is equalised to in-painting an image with known pixel values (shadow scale profile lengths) at original shadow boundary points and unknown values for the remaining of pixels. A spring-metaphor based in-painting method (Bertalmio et al., 2000) is adopted to smoothly interpolate and extrapolate the unknown values.

8.2.2 Synthesisation of Shadow Matte

Our goal is to generate a shadow matte according to user specified shadow properties. Given a shadow mask, whether the original or not, its crisp shadow boundaries can be located using a method for boundary tracing from a binary image (Schalkoff, 1989). The resulting boundaries should exclude the boundaries of the image border. A typical solution to generate a shadow matte is to re-generate the sparse scales of sampling lines perpendicular to shadow boundaries and in-paint for the other unknown shadow scales in image like the algorithm to form a dense shadow scale field from sparse scales described in Chapter 7. However, this can be computationally costly for interactive performance as the in-painting process is comparatively slow and every modification of shadow revokes the in-painting process. Instead, a linear time Euclidean distance transform (Maurer Jr et al., 2003) is used to generate the shadow matte quickly. The procedure for the generation of shadow mattes is described in Algorithm 4 where χ

Algorithm 4: Generation of shadow matte

input : point set of shadow boundary $\mathbf{B}$, field of shadow softness F, synthesised shadow scale attenuation profile A, shadow mask N
output: shadow matte S

Initialise M as a zero matrix in the size of original image;
Initialise S as a matrix of ones in the size of original image;
$M(\mathbf{B}) \leftarrow 1$; `/* mark shadow boundary points */`
$(D, L) \leftarrow \chi(M)$; `/* perform a Euclidean distance transform */`
foreach *point* $p \in \mathbf{B}$ **do**
  $E \leftarrow \{x | L(x) = \phi(p)\}$; `/* find pixels closet to current point */`
  `/* divide the point set` E `into lit and shadow parts */`
  $E_l \leftarrow \{x | x \in E \cap N(x)\}$; `/* lit part */`
  $E_s \leftarrow \{x | x \in E \cap x \notin E_l\}$; `/* shadow part */`
  `/* perform distance conversion */`
  $R \leftarrow D(E)$;
  $R(E_l) = \max(0.5 + D(E_l)/F(p), 1)$;
  $R(E_s) = \min(0.5 - D(E_s)/F(p), 0)$;
  $r \leftarrow 1000$; `/* number of cached intensity attenuation data sites */`
  foreach *RGB colour channel* $c \in \{1, 2, 3\}$ **do** convert distance to shadow scale
    $S(E, c) \leftarrow A(\max(\kappa(rR), 1), c)$;
  end
end

is a function computes Euclidean distance transform (Maurer Jr et al., 2003) (its first output D refers to a matrix of Euclidean distance and its second output L refers to a label index matrix of closest boundary point), ϕ is a function which returns the index of a pixel in image, κ is a function that truncates each value of a set of data.

Algorithm 4 already handles variable shadow shape. To provide more controls for softness, darkness, and colour of shadow, more parameters are added to enable tuning:
Softness of shadow A parameter of global softness adjustment is provided. This is achieved by simply multiplying the field of softness F in Algorithm 4 with a scaling factor v_s as follows:

$$F_n = \max(Fv_s, 2) \tag{8.1}$$

where F_n is the modified field of softness and an additional saturation operation is applied to ensure the width of penumbra is at least 2 pixel wide.
Darkness of shadow A parameter for darkness adjustment is provided. This is achieved by modifying the synthesised shadow scale attenuation profile A in Algorithm 4 using a scaling factor V_d which controls the variation of shadow scale attenuation. The profile A is modified as follows:

$$\begin{cases} m(c) = \mu(A(1 \dots r, c)) \\ A_n(1 \dots r, c) = \max(v_d(A(1 \dots r, c) - m(c)) + m(c), 0) \end{cases} \tag{8.2}$$

where A_n is the updated profile, c is the RGB channel index, r is the number of pre-computed data sites in Algorithm 4, μ is a function which computes the mean of a set of values. An additional saturation operation is applied to ensure the shadow scale is non-negative.
Colour of Shadow Three parameters of colour adjustment are provided. Each of them controls the intensity strength of its RGB colour channel. The 3 parameters are represented as a 1-by-3 vector V_k. The adjustment is done by further modifying Eq. 8.2 as follows:

$$\begin{cases} m(c) = \mu(A(1 \dots r, c)) \\ o_k(c) = A(1, c) \\ d_k(c) = V_k(c) - o_k(c) \\ q_k(c) = (1 - V_k(c))/(1 - o_k(c)) \\ A_n(1..r, c) = \max(q_k(c)v_d(A(1..r, c) - m(c)) + m(c) + d_k(c)/2, 0) \end{cases} \tag{8.3}$$

where o_k is an intensity vector representing the original colour of the shadow, d_k and q_k are the error and ratio between original and adjusted colours of shadow respectively which change the variation and mean of shadow scale attenuation profile A.

After the adjustment, A_n may contain shadow scale values greater than 1, it is thus normalised by dividing all data sites by its lit end value as follows:

$$A_n(1..r, c) = A_n(1..r, c)/A_n(r, c) \tag{8.4}$$

8.2.3 Composition of Shadow

According the Eq. 2.4, the edited shadow can be composited using the Hadamard product of the edited shadow matte and the shadow-free image. This approach of composition ensures that the original properties of the surface background, e.g. texture and reflectance, are preserved.

8.3 User Interaction

In this section, we describe our user interaction and its related algorithm. Fig. 8-2 shows the prototype of our user interface. This interface can be divided into a drawing section and a configuration section.

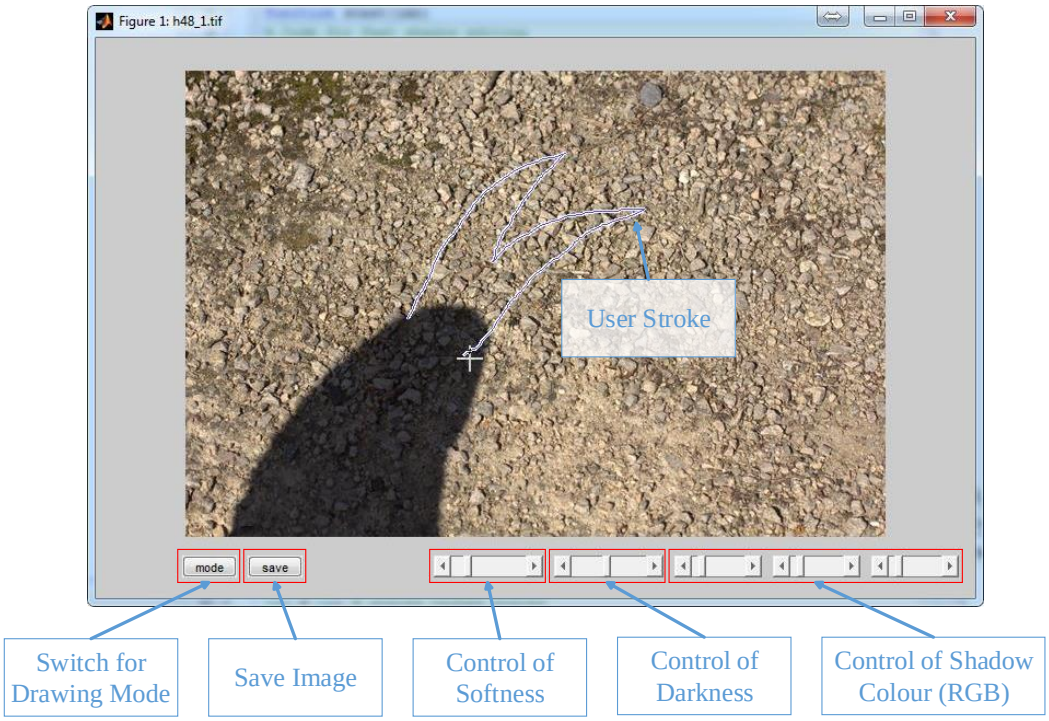


Figure 8-2: *Graphic user interface for shadow editing.*

8.3.1 Drawing Section for Shape Modification

In previous work (Mohan et al., 2007), users are required to specify multiple sparse control points defining shadow boundary. This user interaction requires multiple delicate mouse clicks to ensure accuracy of shadow modification. After the initial specification,

users can drag the previously defined boundary points to change the shape of shadow. However, its modifiable shapes are limited by the initial boundary points which also disallows users to add new segments of shadow. The interaction can be cumbersome for complex shape editing as users have to move multiple boundary points. To alleviate this burden as well as to increase the freedom of editing, we introduce an interaction to modify the shape of shadow by simply adding and subtracting user drawn shapes.

At start-up, our system detects, removes shadow and loads the original shadow according to the original shadow shape. Users then draw the boundary of new shape for addition or subtraction. The system automatically connects the two ends of a boundary curve using a straight line when the boundary is not closed. The drawing actions have two types:

Shadow addition Shadow addition is equivalent to adding or sometimes merging a shape to the original shadow shape. This is done by applying a logical *or* image operation as follows:

$$N_n = N \vee N_a \tag{8.5}$$

where N refers to the shadow mask in Algorithm 4, N_a is the additional shape drawn by user, N_n is the updated shadow mask.

Shadow subtraction Similar to Eq. 8.5, shadow subtraction is done by applying two logical image operations as follows:

$$N_n = N \wedge \neg N_a \tag{8.6}$$

As our shadow matte generation in Algorithm 8.2 already handles arbitrary shadow masks, users can draw any complicated shapes rather than changing a single existing shape in previous work (Mohan et al., 2007).

8.3.2 Configuration Section for Other Controls

The other non-drawing controls are placed in the configuration section. The "mode" toggle button switches drawing mode between addition and subtraction of shape. The parameters for tuning shadow editing model are controlled by scrollbars.

8.4 Evaluation

In this section, we demonstrate our shadow editing system with examples of different types of shadow and show the quantitative evaluation result of our shadow reconstruction based on our state-of-the-art shadow removal data set (Chapter 5).

8.4.1 Demonstration of Shadow Editing

The shadow editing steps for various modification and shadow scenes are show in Fig. 8-3. Please also see our supplementary material for the video demonstration of these examples. Our algorithm is implemented in MATLAB script (non-MEX) and it gives interactive performance on a 2.4G hz machine. Fig. 8-3 also provides a qualitative evaluation of the shadow removal and editing algorithms. The real shadows are modified seamlessly in the images of different scenes. This qualitative demonstration reflects the compatibility of our shadow editing tool. The shadow editing process only changes the illumination and the background features, e.g. texture and reflectance, are preserved.

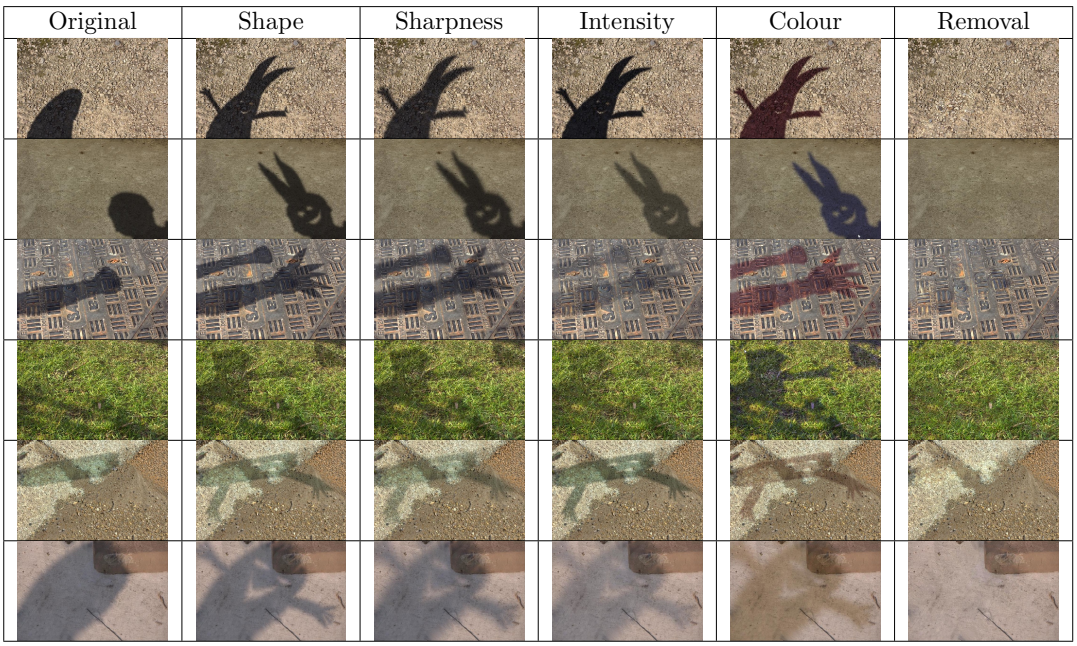


Figure 8-3: *Demonstration of shadow editing in different scenes: The first two rows are examples of easy scenes. The remaining rows are examples for scenes with strong texture background, broken shadow, coloured shadow, and soft shadow respectively.*

8.4.2 Quantitative Evaluation of Shadow Reconstruction

Although our system can produce new shadows perceptually similar to the original shadow, there may still be minor errors introduced in shadow composition phrase. To evaluate any error and judge the quantitative performance of our system, we reconstruct the original shadow image using our shadow editing model with the extract shadow information and compute the Root Mean Square Error (RMSE) between the generated image and the original image. The evaluation is based on our state-of-the-art shadow removal data set (Chapter 5) containing 214 categorised shadow images in variable scenes. As shown in Tab. 8.1, our system produces negligible errors of shadow reconstruction and has similar accuracies for all 4 categories. The accuracy of recon-

struction is often not crucial for shadow editing but can be important for photo forging.

	Texture				Soft				Broken				Colour				Other
	1	2	3	M	1	2	3	M	1	2	3	M	1	2	3	M	
E	1(0)	1(1)	3(1)	2(1)	1(0)	1(1)	3(1)	2(1)	1(0)	2(1)	3(1)	2(1)	1(1)	1(0)	2(1)	2(1)	2(1)
E*	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	1(1)	0(0)	0(0)

Table 8.1: *Error of shadow reconstruction according to 4 attributes. The intensities of image are normalised and the errors shown in this table are in percent (multiplied by 100). The "E" and "E*" indicate the error score where all pixels in the image are used, and just shadow area pixels respectively. For each score of each attribute, the images with other predominant attributes (degree = 3) are not used. Hence, test cases have a strong single bias towards one of the attributes. "Other" refers to a set of shadow cases showing no markedly predominant attributes (degree = 1). "M" refers to the average score for each category. Standard deviations are shown in brackets.*

8.5 Summary

We have presented a shadow editing system with a user-friendly interface. It enables users to freely and quickly modify various properties of existing shadows in images which include shape, darkness, softness, and colour. We have demonstrated and quantitatively evaluated our shadow editing system. Future work includes: 1) the evaluation of the interface with actual users; 2) making use of estimated light sources and geometry from images to provide intelligent suggestions for users to create plausible shadows.

Publications

This chapter is related to the following publication:

- Gong, H. and Cosker, D. (2014a). Interactive shadow editing from single images. In *Workshop on User-Centred Computer Vision, Asian Conference on Computer Vision (ACCV)*

Chapter 9

Conclusions

Shadow removal from natural scenes is still an open problem for scenes with complexity in background texture and properties of shadows, such as softness, colourfulness, and brokenness. In this thesis, we have introduced four complete solutions for user-assisted shadow removal from single images with continuous improvements. These four methods have solved this problem to different degrees. The improvements are generally made in the aspects of user interaction, efficiency and quality of illumination estimation, and amendment of artefacts due to image post-processing. Besides, an algorithm for natural shadow editing from single images is proposed as the application of our shadow removal algorithms. In this chapter, we first summarise our main contributions before describing potential future direction of research in this field.

9.1 Main Contributions

Our contributions can be concluded as follows:

- Texture-preserving shadow removal (Chapter 3): We propose a method which focuses on improving user interaction, the quality of texture preserving penumbra recovery, and plausible umbra recovery.
- Artefact-resistant shadow removal (Chapter 4): This method improves texture-preserving shadow removal by simplifying the optimisation term, removing penumbra residual artefacts by using scale smoothing, and robustly correcting artefacts due to image post-processing using variation alignment.
- First stable and categorised data set for shadow removal evaluation (Chapter 5): This data set overcomes previous limitations in diversity of shadow and consistency between shadow image and its shadow-free ground truth. It contains quantitatively verified high quality ground truth data in variable scene categories along with an open on-line evaluation website to encourage future open comparison.

- Fast shadow removal (Chapter 6): We propose a method that significantly improves the speed and compatibility of variable types of shadow using sub-group illumination change synthesisation.
- Multi-scale shadow removal (Chapter 7): This method improves fast shadow removal by refining smoothness of penumbra recovery on surfaces with varying texture and colour using multi-scale smoothing and further amending artefacts due to image post-processing using multi-scale colour correction.
- Interactive shadow editing (Chapter 8): Based on our shadow removal algorithms, a system for interactive shadow editing from single images is proposed. Its features include the manipulations of shape, distribution, sharpness and darkness of shadows according to the features of existing shadows.

9.2 Future Work

Our evaluation data set contains the shadows in variable scene categories annotated according to four common attributes of shadow which are texture of background, softness, brokenness, colourfulness. The evaluation result reveals that the current state-of-the-art algorithms including ours still have not completely resolved the problem of removing these popular types of shadows. It is found that coloured shadows are significantly difficult to remove than shadows of the other types. This is due to the ambiguity in recognising the background texture when the shadow is non-grey. Besides the chances for improving shadow removal quality for the categorised popular shadows in our data set, there are opportunities for exploring the removal for highly-complicated shadows caused by the following factors:

- Complicated light source: The light sources can be multiple and in different colours. This results in overlapping shadows with different colours which adds more ambiguities in recognising illumination changes and the complexity of shadow geometry.
- Complicated occluder: When the occluders are transparent, some light may pass through the occluders and some complicated refractions occur. This results in highly varying illumination in umbra area.
- Under-saturation: When the shadow area is too dark and the exposure time is not long enough, the intensities in the shadow are near zero and the noise in the shadow is also relatively high. It is thus impossible to recover the shadow area by simple relighting. In-painting and de-noising techniques may be required to cope with this situation.

Although automatic shadow detection is possible, reliability has always been a problem, future work would also be improving the reliability of automatic shadow detection.

Appendix A

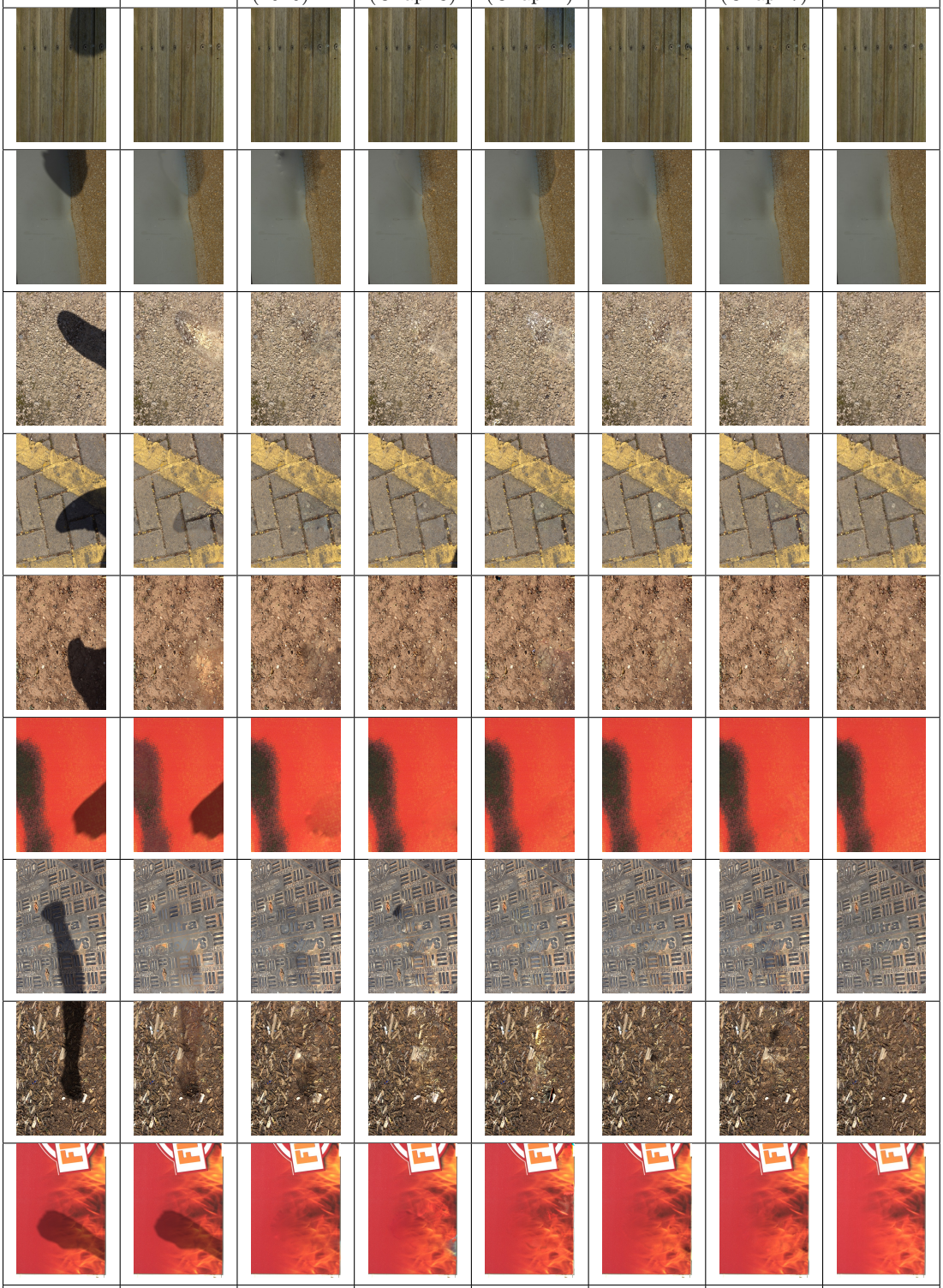
Supplementary Visual Results of Shadow Removal

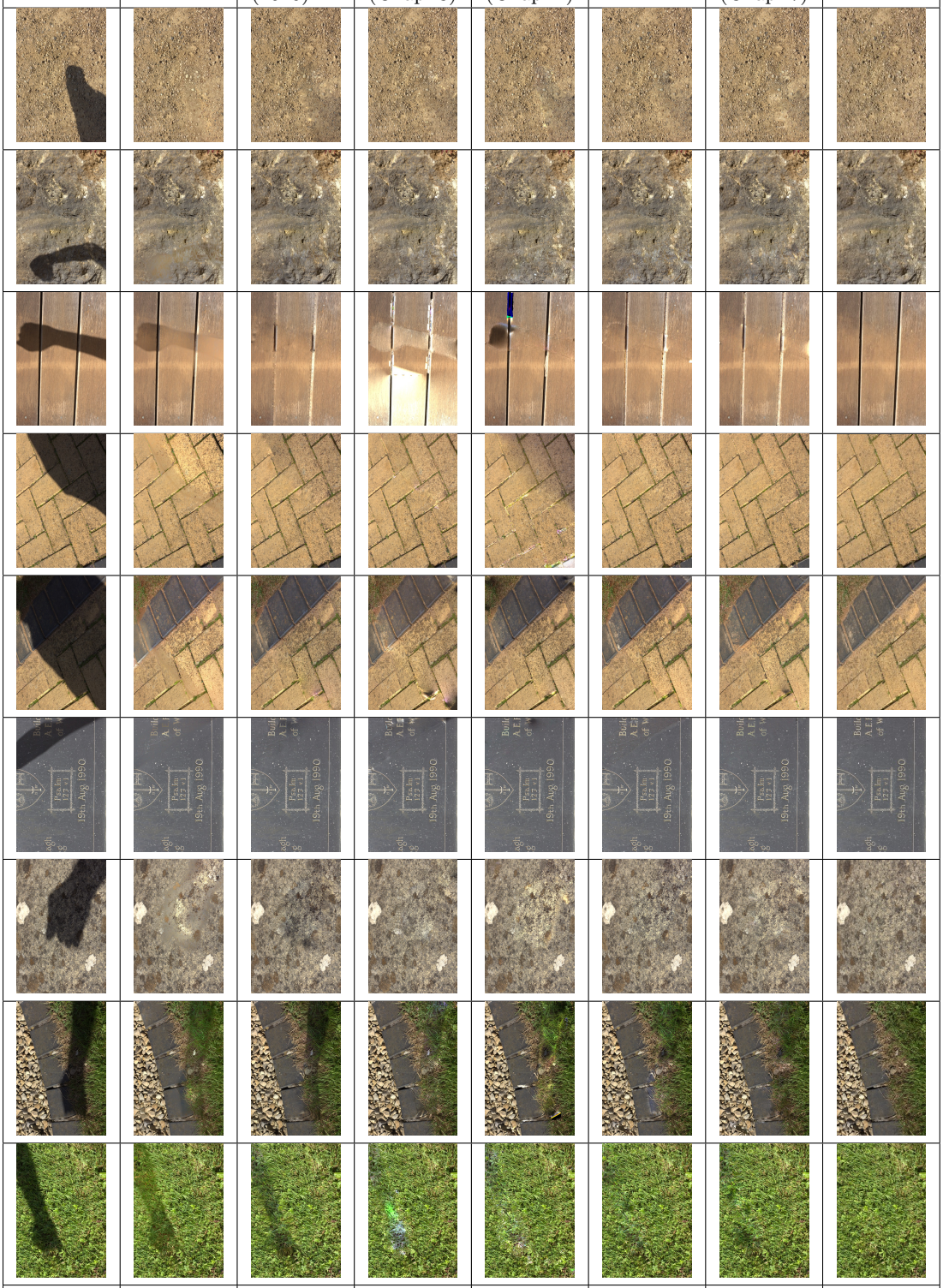
This chapter shows the extended visual comparisons of our shadow removal results. The results of two state-of-the-art shadow removal methods (Guo et al., 2012; Su and Chen, 2010) are generated by using their original implementations. We decide to omit the results of (Yang et al., 2012) as its quantitative scores are relatively much lower and thus less interesting to compare than the others. The representative test results are shown in the tables in the following pages. The presented images are displayed in higher resolution. The images of shadow-free ground truth are displayed in lower resolution for preventing the abuse of our shadow removal on-line benchmark. For the best visual examination, please zoom in the electronic document.

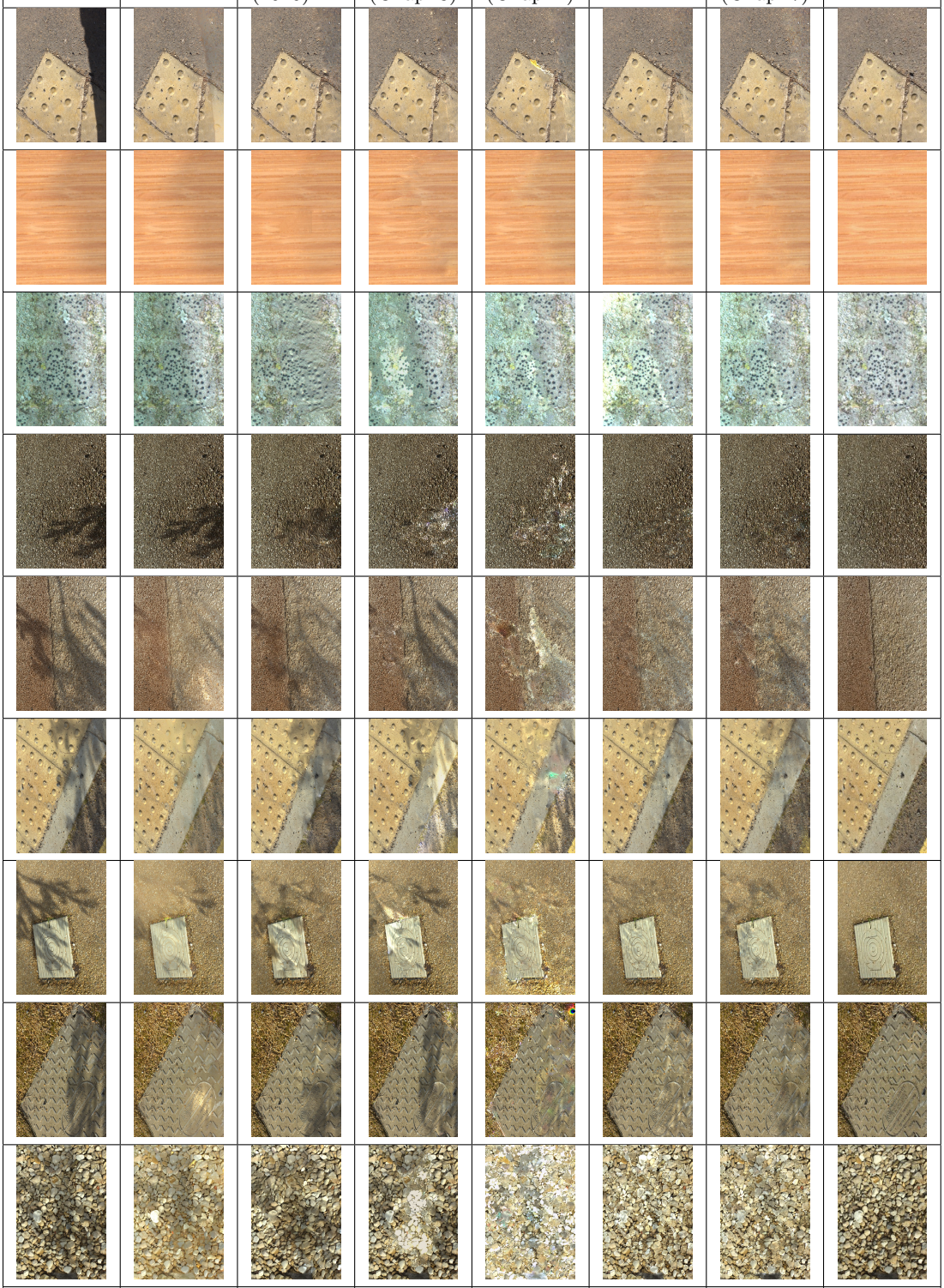
Original	Guo et al. (2012)	Su and Chen (2010)	Texture-Preserving (Chap. 3)	Artefact-Resistant (Chap. 4)	Fast (Chap. 6)	Multi-Scale (Chap. 7)	Ground Truth
							
Original	Guo et al. (2012)	Su and Chen (2010)	Texture-Preserving (Chap. 3)	Artefact-Resistant (Chap. 4)	Fast (Chap. 6)	Multi-Scale (Chap. 7)	Ground Truth

Original	Guo et al. (2012)	Su and Chen (2010)	Texture-Preserving (Chap. 3)	Artefact-Resistant (Chap. 4)	Fast (Chap. 6)	Multi-Scale (Chap. 7)	Ground Truth
							
Original	Guo et al. (2012)	Su and Chen (2010)	Texture-Preserving (Chap. 3)	Artefact-Resistant (Chap. 4)	Fast (Chap. 6)	Multi-Scale (Chap. 7)	Ground Truth

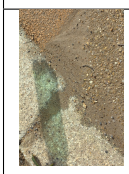
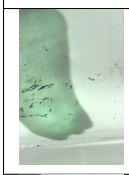
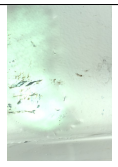
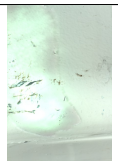
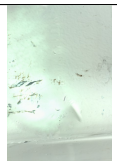
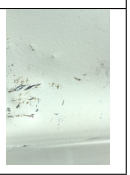
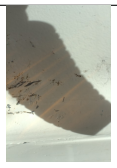
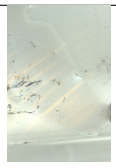
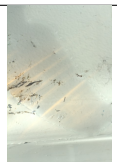
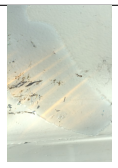
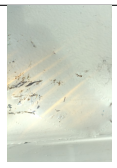
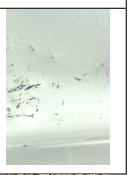
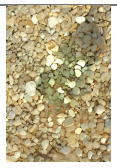
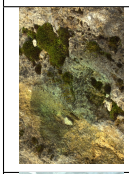
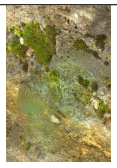
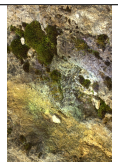
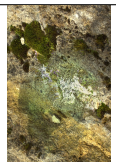
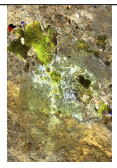
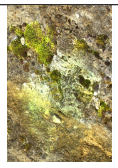
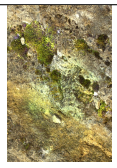
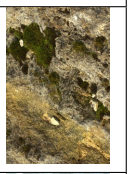
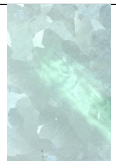
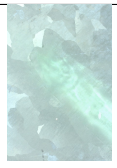
Original	Guo et al. (2012)	Su and Chen (2010)	Texture-Preserving (Chap. 3)	Artefact-Resistant (Chap. 4)	Fast (Chap. 6)	Multi-Scale (Chap. 7)	Ground Truth
							
Original	Guo et al. (2012)	Su and Chen (2010)	Texture-Preserving (Chap. 3)	Artefact-Resistant (Chap. 4)	Fast (Chap. 6)	Multi-Scale (Chap. 7)	Ground Truth

Original	Guo et al. (2012)	Su and Chen (2010)	Texture-Preserving (Chap. 3)	Artefact-Resistant (Chap. 4)	Fast (Chap. 6)	Multi-Scale (Chap. 7)	Ground Truth
							
Original	Guo et al. (2012)	Su and Chen (2010)	Texture-Preserving (Chap. 3)	Artefact-Resistant (Chap. 4)	Fast (Chap. 6)	Multi-Scale (Chap. 7)	Ground Truth

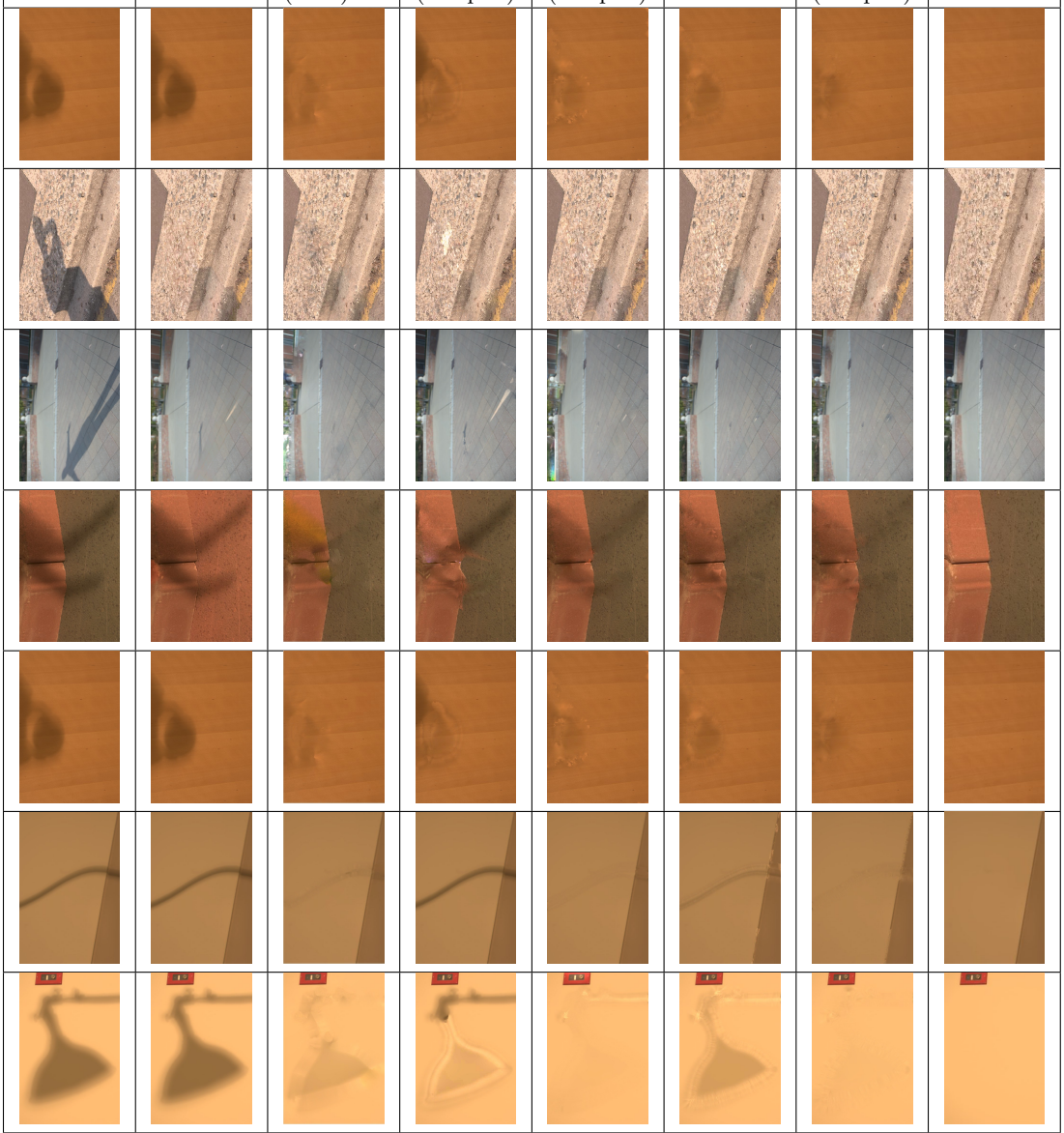
Original	Guo et al. (2012)	Su and Chen (2010)	Texture-Preserving (Chap. 3)	Artefact-Resistant (Chap. 4)	Fast (Chap. 6)	Multi-Scale (Chap. 7)	Ground Truth
							
Original	Guo et al. (2012)	Su and Chen (2010)	Texture-Preserving (Chap. 3)	Artefact-Resistant (Chap. 4)	Fast (Chap. 6)	Multi-Scale (Chap. 7)	Ground Truth

Original	Guo et al. (2012)	Su and Chen (2010)	Texture-Preserving (Chap. 3)	Artefact-Resistant (Chap. 4)	Fast (Chap. 6)	Multi-Scale (Chap. 7)	Ground Truth
							
Original	Guo et al. (2012)	Su and Chen (2010)	Texture-Preserving (Chap. 3)	Artefact-Resistant (Chap. 4)	Fast (Chap. 6)	Multi-Scale (Chap. 7)	Ground Truth

Original	Guo et al. (2012)	Su and Chen (2010)	Texture-Preserving (Chap. 3)	Artefact-Resistant (Chap. 4)	Fast (Chap. 6)	Multi-Scale (Chap. 7)	Ground Truth
							
Original	Guo et al. (2012)	Su and Chen (2010)	Texture-Preserving (Chap. 3)	Artefact-Resistant (Chap. 4)	Fast (Chap. 6)	Multi-Scale (Chap. 7)	Ground Truth

Original	Guo et al. (2012)	Su and Chen (2010)	Texture-Preserving (Chap. 3)	Artefact-Resistant (Chap. 4)	Fast (Chap. 6)	Multi-Scale (Chap. 7)	Ground Truth
							
							
							
							
							
							
							
							
							
Original	Guo et al. (2012)	Su and Chen (2010)	Texture-Preserving (Chap. 3)	Artefact-Resistant (Chap. 4)	Fast (Chap. 6)	Multi-Scale (Chap. 7)	Ground Truth

Original	Guo et al. (2012)	Su and Chen (2010)	Texture-Preserving (Chap. 3)	Artefact-Resistant (Chap. 4)	Fast (Chap. 6)	Multi-Scale (Chap. 7)	Ground Truth
							
Original	Guo et al. (2012)	Su and Chen (2010)	Texture-Preserving (Chap. 3)	Artefact-Resistant (Chap. 4)	Fast (Chap. 6)	Multi-Scale (Chap. 7)	Ground Truth

Original	Guo et al. (2012)	Su and Chen (2010)	Texture-Preserving (Chap. 3)	Artefact-Resistant (Chap. 4)	Fast (Chap. 6)	Multi-Scale (Chap. 7)	Ground Truth
							
Original	Guo et al. (2012)	Su and Chen (2010)	Texture-Preserving (Chap. 3)	Artefact-Resistant (Chap. 4)	Fast (Chap. 6)	Multi-Scale (Chap. 7)	Ground Truth

Bibliography

Ahmed, N., Natarajan, T., and Rao, K. R. (1974). Discrete cosine transform. *Computers, IEEE Transactions on*, 100(1):90–93.

Al-Najdawi, N., Bez, H. E., Singhai, J., and Edirisinghe, E. (2012). A survey of cast shadow detection algorithms. *Pattern Recognition Letters*, 33(6):752 – 764.

Arbel, E. and Hel-Or, H. (2007). Texture-preserving shadow removal in color images containing curved surfaces. In *CVPR*, volume 7, pages 1–8.

Arbel, E. and Hel-Or, H. (2011). Shadow removal using intensity surfaces and texture anchor points. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 33(6):1202–1216.

Barron, J. T. and Malik, J. (2012). Shape, albedo, and illumination from a single image of an unknown object. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 334–341. IEEE.

Barron, J. T. and Malik, J. (2013). Intrinsic scene properties from a single rgb-d image. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 17–24. IEEE.

Bertalmio, M., Sapiro, G., Caselles, V., and Ballester, C. (2000). Image inpainting. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '00, pages 417–424.

Bousseau, A., Paris, S., and Durand, F. (2009). User-assisted intrinsic images. In *ACM Transactions on Graphics (TOG)*, volume 28, page 130. ACM.

Cavallaro, A., Salvador, E., and Ebrahimi, T. (2005). Shadow-aware object-based video processing. In *Vision, Image and Signal Processing, IEE Proceedings-*, volume 152, pages 398–406. IET.

Center, S. I. A. I. and Barrow, H. (1978). *Recovering intrinsic scene characteristics from images*. Technical note. Artificial Intelligence Center, SRI International.

Chang, C.-J., Hu, W.-F., Hsieh, J.-W., and Chen, Y.-S. (2002). Shadow elimination for effective moving object detection with gaussian models. In *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, volume 2, pages 540 – 543 vol.2.

Chen, C.-T., Su, C.-Y., and Kao, W.-C. (2010). An enhanced segmentation on vision-based shadow removal for vehicle detection. In *Green Circuits and Systems (ICGCS), 2010 International Conference on*, pages 679 –682.

Chen, Q. and Koltun, V. (2013). A simple model for intrinsic image decomposition with depth cues. In *Computer Vision (ICCV), 2013 IEEE International Conference on*, pages 241–248. IEEE.

Cliffs, E. (1990). *Two-Dimensional Signal and Image Processing*. Prentice Hall.

Comaniciu, D. and Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. *PAMI*, 24(5):603–619.

Criminisi, A., Pérez, P., and Toyama, K. (2004). Region filling and object removal by exemplar-based image inpainting. *Image Processing, IEEE Transactions on*, 13(9):1200–1212.

Cucchiara, R., Grana, C., Piccardi, M., and Prati, A. (2003). Detecting moving objects, ghosts, and shadows in video streams. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(10):1337–1342.

Drew, M., Lu, C., and Finlayson, G. (2006). Removing shadows using flash/noflash image edges. In *IEEE International Conference on Multimedia and Expo (ICME)*, pages 257–260.

Duchon, J. (1977). Splines minimizing rotation-invariant semi-norms in sobolev spaces. In *Constructive theory of functions of several variables*, pages 85–100. Springer.

Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *KDD*, volume 96, pages 226–231.

Finlayson, G., Fredembach, C., and Drew, M. S. (2007). Detecting illumination in images. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8. IEEE.

Finlayson, G. D., Drew, M. S., and Lu, C. (2009). Entropy minimization for shadow removal. *IJCV*, 85(1):35–57.

Finlayson, G. D., Hordley, S. D., and Drew, M. S. (2002a). Removing shadows from images. In *ECCV*, pages 823–836. Springer.

Finlayson, G. D., Hordley, S. D., and Drew, M. S. (2002b). Removing shadows from images using retinex. In *Color Imaging Conference*, pages 73–79.

Finlayson, G. D., Hordley, S. D., Lu, C., and Drew, M. S. (2006). On the removal of shadows from images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 28(1):59–68.

Frankot, R. T. and Chellappa, R. (1988). A method for enforcing integrability in shape from shading algorithms. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 10(4):439–451.

Fredembach, C. and Finlayson, G. (2006). Simple shadow removal. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 1, pages 832–835. IEEE.

Fredembach, C. and Finlayson, G. D. (2005). Hamiltonian path based shadow removal. In *BMVC*, pages 502–511.

Fredembach, C. and Süsstrunk, S. (2010). Automatic and accurate shadow detection from (potentially) a single image using near-infrared information. Technical report, Institute of Electrical and Electronics Engineers.

Friedman, J. H., Bentley, J. L., and Finkel, R. A. (1977). An algorithm for finding best matches in logarithmic expected time. *ACM Transactions on Mathematical Software (TOMS)*, 3(3):209–226.

Fritsch, F. N. and Carlson, R. E. (1980). Monotone piecewise cubic interpolation. *SIAM Journal on Numerical Analysis*, 17(2):238–246.

Garces, E., Munoz, A., Lopez-Moreno, J., and Gutierrez, D. (2012). Intrinsic images by clustering. In *Computer Graphics Forum*, volume 31, pages 1415–1424. Wiley Online Library.

Garcia, D. (2010). Robust smoothing of gridded data in one and higher dimensions with missing values. *Computational Statistics & Data Analysis*, 54(4):1167–1178.

Gong, H. and Cosker, D. (2014a). Interactive shadow editing from single images. In *Workshop on User-Centred Computer Vision, Asian Conference on Computer Vision (ACCV)*.

Gong, H. and Cosker, D. (2014b). Interactive shadow removal and ground truth for variable scene categories. *International Journal of Computer Vision (IJCV)*. under revision.

Gong, H. and Cosker, D. (2014c). Interactive shadow removal and ground truth for variable scene categories. In *British Machine Vision Conference (BMVC)*. Best Student Paper.

Gong, H. and Cosker, D. (2014d). User-assisted image shadow removal. *Computer Vision and Image Understanding (CVIU)*. under review.

Gong, H., Cosker, D., Li, C., and Brown, M. (2013). User-aided single image shadow removal. In *IEEE Proc. International Conference on Multimedia and Expo (ICME)*.

Guo, R., Dai, Q., and Hoiem, D. (2012). Paired regions for shadow detection and removal. *PAMI*, PP(99):1–1.

Huang, J.-B. and Chen, C.-S. (2009). Moving cast shadow detection using physics-based features. In *CVPR*, pages 2310 –2317.

Huang, X., Hua, G., Tumblin, J., and Williams, L. (2011). What characterizes a shadow boundary under the sun and sky? In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 898–905.

Huffman, D. A. et al. (1952). A method for the construction of minimum redundancy codes. *proc. IRE*, 40(9):1098–1101.

J. J. Yoon, C. and T.J.Ellis (2002). Shadowflash: an approach for shadow removal in an active illumination environment. In *BMVC*.

Katramados, I., Crumpler, S., and Breckon, T. (2009). Real-time traversable surface detection by colour space fusion and temporal analysis. In *Proc. Int. Conf. on Computer Vision Systems*, volume 5815, pages 265–274.

Lalonde, J.-F., Efros, A. A., and Narasimhan, S. G. (2010). Detecting ground shadows in outdoor consumer photographs. In *ECCV*, pages 322–335.

Land, E. H. (1971). The retinex theory of color vision. *J Opt Soc Am*, 61:1–11.

Lee, K., Zhao, Q., Tong, X., Gong, M., Izadi, S., Lee, S., Tan, P., and Lin, S. (2012). Estimation of intrinsic image sequences from image+depth video. In *ECCV 2012*, volume 7577, pages 327–340. Springer Berlin Heidelberg.

Levin, A., Lischinski, D., and Weiss, Y. (2008). A closed-form solution to natural image matting. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(2):228–242.

Liu, F. and Gleicher, M. (2008). Texture-consistent shadow removal. In *ECCV*, pages 437–450.

Martin, D. R., Fowlkes, C. C., and Malik, J. (2004). Learning to detect natural image boundaries using local brightness, color, and texture cues. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(5):530–549.

Maurer Jr, C. R., Qi, R., and Raghavan, V. (2003). A linear time algorithm for computing exact euclidean distance transforms of binary images in arbitrary dimensions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(2):265–270.

Mitchell, T. M. (1997). Machine learning. 1997. *Burr Ridge, IL: McGraw Hill*, 45.

Mohan, A., Tumblin, J., and Choudhury, P. (2007). Editing soft shadows in a digital photograph. *IEEE Computer Graphics and Applications*, 27(2):23–31.

Nascimento, S., Ferreira, F. P., and Foster, D. H. (2002). Statistics of spatial cone-excitation ratios in natural scenes. *JOSA A*, 19(8):1484–1490.

Nocedal, J. and Wright, S. (2006). Numerical optimization. Springer series in operations research, chapter 18. Springer, second edition.

Paris, S. and Durand, F. (2009). A fast approximation of the bilateral filter using a signal processing approach. *IJCV*, 81(1):24–52.

Pennebaker, W. B. and Mitchell, J. L. (1993). *JPEG: Still image data compression standard*. Springer.

Poynton, C. (2012). *Digital video and HD: Algorithms and Interfaces*. Elsevier.

Reinhard, E., Ashikhmin, M., Gooch, B., and Shirley, P. (2001). Color transfer between images. *IEEE Computer Graphics and Applications*, 21(5):34–41.

Rhemann, C., Rother, C., Wang, J., Gelautz, M., Kohli, P., and Rott, P. (2009). A perceptually motivated online benchmark for image matting. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 1826–1833. IEEE.

Rother, C., Kiefel, M., Zhang, L., Schölkopf, B., and Gehler, P. V. (2011). Recovering intrinsic images with a global sparsity prior on reflectance. In *Advances in neural information processing systems*, pages 765–773.

Salamati, N., Germain, A., and Susstrunk, S. (2011). Removing shadows from images using color and near-infrared. In *IEEE International Conference on Image Processing (ICIP)*, pages 1713 –1716.

Salvador, E., Cavallaro, A., and Ebrahimi, T. (2004). Cast shadow segmentation using invariant color features. *Computer vision and image understanding*, 95(2):238–259.

Schalkoff, R. J. (1989). *Digital image processing and computer vision*, volume 286. Wiley New York.

Seber, G. A. (2009). *Multivariate observations*, volume 252. John Wiley & Sons.

Serra, M., Penacchio, O., Benavente, R., and Vanrell, M. (2012). Names and shades of color for intrinsic image estimation. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 278–285. IEEE.

Shen, J., Yang, X., Jia, Y., and Li, X. (2011). Intrinsic images using optimization. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 3481–3487. IEEE.

Shen, L. and Yeo, C. (2011). Intrinsic images decomposition using a local and global sparse representation of reflectance. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 697–704. IEEE.

Shor, Y. and Lischinski, D. (2008). The shadow meets the mask: Pyramid-based shadow removal. *Comput. Graph. Forum*, 27(2):577–586.

Su, Y.-F. and Chen, H. H. (2010). A three-stage approach to shadow field estimation from partial boundary information. *IEEE Trans. on Image Processing*, 19(10):2749–2760.

Whitaker, R. T. (1998). A level-set approach to 3d reconstruction from range data. *International Journal of Computer Vision*, 29(3):203–231.

Witten, I. H., Neal, R. M., and Cleary, J. G. (1987). Arithmetic coding for data compression. *Communications of the ACM*, 30(6):520–540.

Wu, T.-P. and Tang, C.-K. (2005). A bayesian approach for shadow extraction from a single image. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 1, pages 480–487. IEEE.

Wu, T.-P., Tang, C.-K., Brown, M. S., and Shum, H.-Y. (2007). Natural shadow matting. *ACM Transactions on Graphics (TOG)*, 26(2):8.

Xiao, Y., Tsougenis, E., and Tang, C.-K. (2014). Shadow removal from single rgb-d images. In *Computer Vision and Pattern Recognition, 2014. CVPR 2014. IEEE Conference on*. IEEE.

Yang, Q., Tan, K.-H., and Ahuja, N. (2012). Shadow removal using bilateral filtering. *IEEE Trans. on Image Proc.*, 21(10):4361–4368.

Yoneyama, A., Yeh, C., and Kuo, C.-C. (2003). Moving cast shadow elimination for robust vehicle extraction based on 2d joint vehicle/shadow models. In *Proceedings. IEEE Conference on Advanced Video and Signal Based Surveillance, 2003.*, pages 229–236.

Yoon, J., Koch, C., and Ellis, T. J. (2002). Shadowflash: an approach for shadow removal in an active illumination environment. In *BMVC*, pages 1–10.

Zhu, J., Samuel, K. G. G., Masood, S., and Tappen, M. F. (2010). Learning to recognize shadows in monochromatic natural images. In *CVPR*.